

Digitized by the Internet Archive
in 2024 with funding from
University of Alberta Library

https://archive.org/details/ajer_1975



The Alberta Journal of Educational Research

Volume XXI, No. 1

March, 1975

CONTENTS

The judgment of simple deductive arguments by pre-service elementary school teachers	1
<i>L. C. Jansson</i>	
Validity of Expectation—A Developmental Function	11
<i>M. M. Clifford</i>	
The Interview Reassessed	18
<i>L. Foster and M. Nixon</i>	
Does the Dolch Do?	23
<i>V. Froese</i>	
Studying Problem Solving Behavior in Young Children—Some Methodological Considerations	28
<i>D. Nelson</i>	
Controlled vs. Natural Setting—Some Implications for Behavioral Analysis and Change in Classroom Situations	39
<i>J. Martin</i>	
Day Care Teacher Practices and Beliefs	46
<i>S. Verma and D. L. Peters</i>	
Learning the Language of Reading—An Exploratory Study	56
<i>A. D. Forester</i>	
The Works of Jean Piaget Published in English	63
<i>B. M. Blackall</i>	

FACULTY OF EDUCATION
The University of Alberta

L. C. JANSSEN

University of Manitoba

The Judgment of Simple Deductive Arguments by Pre-Service Elementary School Teachers

The purpose was the assessment of pre-service elementary school teachers' abilities to judge selected deductive arguments. 140 Ss were given tests requiring validity judgments of five varieties of argument in isomorphic class and conditional forms. The judgment of arguments of the class type were found to be significantly easier than the judgment of arguments of the conditional type. A significant difference was also found among abilities to judge the validity of the five varieties (principles) of argument. Furthermore, it was possible to rank these five principles in order of difficulty, and the ranking was constant over the two types of reasoning. A significant interaction of type and principle was found. The percentage of Ss mastering each of the five principles was also determined. (Dr. Jansson is Associate Professor in the Department of Curriculum, Faculty of Education, The University of Manitoba.)

Over the past decade a number of researchers, e.g., Hill (1961), Howell (1965), O'Brien and Shapiro (1968), O'Brien, Shapiro, and Reali (1971), have examined the developmental aspects of children's abilities to handle deductive arguments, valid and invalid, of simple sentential and class types. These studies have been concerned with logical skills *per se*.

In addition to the studies cited above, there is a set of investigations spawned by the work of the Cornell Critical Thinking Project. Some of these (Gardiner, 1966; Paulus, 1967; Roberge, 1970) have, like the above studies, looked at developmental patterns in logical skills development, while others (Ennis, 1959 and 1962; Jansson, 1970) have focussed on the broader notion of critical thinking ability, an important component of which is the ability to assess the validity of logical arguments.

Hill (1961) suggested that children of ages six, seven, and eight "are able to deal very effectively with verbal premises that call for hypothetical reasoning . . .," to which O'Brien and Shapiro added some important qualifications based upon their more open response form. The latter study

used the same item types (sentential logic, classical syllogism, and logic of quantification) as Hill, but included a "not enough clues" distractor, and concluded that the abilities found by Hill "cannot be taken for granted in children of this age." Neither of these studies involved explicit instruction in the logical abilities. Roberge and Paulus (1971) also used an assessment form in their investigation of children's abilities in class and conditional reasoning. Overall in grades 4, 6, 8, and 10 they found a significant ($p < .05$) difference in children's abilities to handle the two types of arguments. However, significant differences did not appear at each grade level tested, and in fact, the means were reversed at grade 8 in favor of the conditional, rather than class reasoning. The complete analysis showed a strong interaction between type of reasoning and a three-level content dimension (concrete-familiar, abstract, and symbolic). As these researchers state, the "results suggest that it is psychologically and pedagogically feasible to introduce brief units on class and conditional reasoning as early as the fourth grade." They also indicate possible sources of difficulty (e.g., content) toward which teaching strategies could be directed. "Such information should be valuable to mathematics teachers, curriculum planners, and psychologists alike (p. 199)."

The Present Study

The focus in the present study is on the determination of an individual's ability to assess a logical argument, rather than on his ability to deduce. It is pointed out by Paulus (1967) that these abilities are different (deducing is more difficult) in certain important respects, but that the difference seems to decrease as the students move from grade 5 to grade 11, the range covered in his study. Again, there are important interactions with other variables such as content and form, but this is not of immediate concern here.

All of the studies cited here, while they have many points of disagreement, appear to suggest that some instruction in logical reasoning may be appropriate by at least grade 4, if not before. In fact, such materials is finding its way into elementary mathematics textbook series. This has important implications for our teacher training programs.

The above discussion and the cited studies themselves give some insight into the status of children's logical reasoning abilities. The present study was formulated as a status study of pre-service elementary teachers' abilities to judge logical arguments employing five specified principles of reasoning in both class and conditional forms. Clearly teachers must possess these skills if they are going to provide explicit (or even implicit) instruction in them.

More specifically, answers were sought to the following questions:

1. Do the teachers' abilities to judge deductive arguments vary according to the type of reasoning?
2. Do the teachers' abilities to judge deductive arguments vary according to the principle of reasoning?
3. Do the teachers' abilities to judge deductive arguments in class and conditional form vary according to the principle?
4. What proportion of the teachers have mastered each of the five principles of logical reasoning?

Judgment of Simple Deductive Arguments

In each of the above questions “teachers” refers to pre-service elementary school teachers taking a regular elementary education mathematics methods course. Each subject has had at least one course in mathematics which may have included some material on formal logic. The answers to the questions should prove valuable in structuring a unit on logic for such students and in providing instructional hypotheses relative to their learning of logical skills.

Method

Tests

The tests used in this investigation were:

1. *The Roberge Class Reasoning Test*, Form AN
2. *The Roberge Conditional Reasoning Test*, Form BN

Both were developed by Roberge and used with permission. These instruments have evolved from earlier ones used by Roberge and Paulus and stem from the conceptual and empirical work of the Cornell Critical Thinking Project.

The test items evaluated the subjects’ abilities on five principles of deductive reasoning. Item format is as follows:

Suppose you know that

premise

premise

Then would this be true?

Conclusion

Three responses were possible and were defined as follows for the examinees:

- | | |
|----------|---|
| 1. YES | It must be true. |
| 2. NO | It can’t be true. |
| 3. MAYBE | It may be true or it may not be true. You weren’t told enough to be <i>certain</i> whether it is “YES” or “NO.” |

Principles

The five principles of reasoning of each type are exhibited in Tables 1 and 2. More background information on these and related tests may be found in Paulus (1967), Roberge (1970), and Roberge and Paulus (1971).

Procedures

The sample of 140 subjects was chosen randomly from approximately 200 junior and senior undergraduates enrolled in an elementary mathematics methods course during the fall of 1972. The tests were administered during the first two weeks of the term, one week apart. Half of the group received the Conditional Reasoning test first, the other half took the Class Reasoning test first.

In order to answer the questions posed, a 2x2x5 analysis of variance (fixed factors) with repeated measures on the last two factors was performed (Winer, 1962, 319-337). The first factor was that of test sequence. The second and third factors, the ones of critical interest, were, respectively, types of reasoning, and principles. The Newman-Keuls method was used for *post hoc* comparisons of means. Further analysis of the data indicated the proportion mastering each of the principles.

TABLE 1
BASIC PRINCIPLES OF CLASS REASONING

Principle	Validity	Symbolic Form	Concrete Form
1	YES	All A's are B's x is an A ∴ x is a B	All of the dogs that are brown are named Rover. The dog is brown. ∴ The dog's name is Rover.
2	NO	All A's are B's x is a B ∴ x is an A	All of the paintings that belong to Don are paintings of horses. This is a painting of a horse. ∴ This painting belongs to Don.
3	NO	All A's are B's x is not an A ∴ x is not a B	All of the cars in the garage that are Mr. Smith's are black. The car in the garage is not Mr. Smith's. ∴ The car is not black.
4	YES	All A's are B's x is not a B ∴ x is not an A	All of the cats that are not named Tabby are also not white. The cat is white. ∴ The cat's name is Tabby.
5	YES	All A's are B's All B's are C's ∴ All A's are C's	All of the coats in the closet that belong to Donna are blue. All of the coats that are blue have white buttons. ∴ All of the coats in the closet that belong to Donna have white buttons.

Results and discussion

The dependent variable in the analysis of variance was subjects' scores on the five principle subtests. Since each subtest contained eight items, possible scores ranged from 0 to 8. The analysis of variance is presented in Table 3. Tables 4, 5 and 6 present means and differences between means while Table 7 is a summary of the proportions of subjects mastering each principle.

The *F* ratios for the factor test sequence (Table 3) and its interactions with the factors types of reasoning and principles were non-significant ($p > .05$).

The significant ($p < .05$) *F* ratio for the factor types of reasoning indicated a significantly higher mean in class reasoning than in conditional reasoning when considered together with the total mean scores for each type (Table 4).

The significant ($p < .001$) *F* ratio for the factor principles of reasoning indicates that abilities varied significantly among the five principles of reasoning. The means given in the rows of Table 4 entitled "combined" refer to means of all individuals over each principle, weighted proportionally to standard deviation, *i.e.*, scores are added and divided by two. The

Judgment of Simple Deductive Arguments

TABLE 2
BASIC PRINCIPALS OF CONDITIONAL REASONING

Principle	Validity	Symbolic Form	Concrete Form
1	YES	if p, then q $\frac{p}{q}$	If the dog is brown, then his name is Rover. The dog is brown. The dog's name is Rover.
2	NO	if p, then q $\frac{q}{p}$	If the painting belongs to Don, then it is a painting of a horse. This is a painting of a horse. This painting belongs to Don.
3	NO	if p, then q $\frac{\text{not } p}{\text{not } q}$	If the car in the garage is Mr. Smith's, then it is black. The car in the garage is not Mr. Smith's. The car is not black.
4	YES	if p, then q $\frac{\text{not } q}{\text{not } p}$	If the cat's name is not Tabby, then she is not white. The cat is white. The cat's name is Tabby.
5	YES	if p, then q if q, then r $\frac{\text{if p, then r}}$	If the coat in the closet belongs to Donna, then it is blue. If it is blue, then it has white buttons. If the coat in the closet belongs to Donna, then it has white buttons.

TABLE 3
ANALYSIS OF VARIANCE FOR TEST SEQUENCE, TYPES OF REASONING
AND PRINCIPLES OF REASONING

Source	df	MS	F
<u>Between Subjects</u>			
A (Test Sequence)	1	0.93	0.10
Subjects Within Groups	138	9.49	
<u>Within Subjects</u>			
J (Types of Reasoning)	1	10.29	4.97*
A x J	1	7.73	3.74
J x Subjects Within Groups	138	2.07	
K (Principles of Reasoning)	4	813.23	172.341**
A x K	4	2.34	0.496
K x Subjects Within Groups	552	4.72	
J x K	4	13.77	9.170**
A x J x K	4	8.22	5.476**
J x K x Subjects Within Groups	552	1.50	

* $p < .05$
** $p < .001$

TABLE 4
CROSS TABULATION OF MEANS FOR ANALYSIS OF VARIANCE

Test Sequence	Type of Reasoning	Principle					
		1	2	3	4	5	All
Conditional - Class	Conditional	6.67	2.94	3.59	4.60	6.14	23.94
	Class	7.35	3.40	3.94	4.59	6.26	25.54
	Combined	7.01	3.17	3.76	4.59	6.20	24.74
Class - Conditional	Conditional	7.36	2.30	3.47	4.76	6.54	24.43
	Class	7.07	3.50	3.70	4.47	5.80	24.54
	Combined	7.21	2.90	3.59	4.61	6.17	24.49
Both	Conditional	7.01	2.62	3.53	4.68	6.34	24.19
	Class	7.21	3.45	3.82	4.53	6.03	25.04
	Combined	7.11	3.04	3.68	4.60	6.19	24.62

analysis of pairwise comparisons (Table 5) indicated that all are significant and that thus the principles can be ranked in order of difficulty: 2,3,4,5,1. This ranking applies both to the combined scores and within each type. Although this order is somewhat different than that found by Roberge (1970) for tenth graders: 3,2,4,5,1, the reason for the shift is not clear. Differences between means for the remaining types X principles simple main affects comparisons are presented in Table 6. Although few unexpected patterns emerge clearly, three characteristics of the matrix are noteworthy. First, an examination of the main diagonal reveals significant differences between class and conditional reasoning types for both invalid principles, affirming the consequent and denying the antecedent. Both principles appear more difficult to judge in the conditional reasoning form. Secondly, the familiar transitivity property appeared easier in the conditional form. Finally, the judgment of a *modus ponens* form of argument of the class type was significantly easier than all other arguments except *modus ponens* itself in the conditional form.

A further analysis was made to determine the proportion of subjects who had mastered each of the five principles. These results are presented in Table 7. Each principle subtest contained eight items and mastery was defined as 6/8. Using cumulative binomial probabilities, it can be determined that the probability of showing mastery of a single principle of one type by chance alone is approximately .02. The fourth column of Table 7, headed "Both" gives the proportion who mastered the principle on both class and conditional forms.

It is instructive to compare the results obtained in this study with those of previous studies. In a recent summary of research related to children's comprehension of deductive reasoning schemes Roberge (1972) presented a table showing percentages of students who had mastered each of the prin-

TABLE 5
DIFFERENCES FOR NEWMAN-KEULS MAIN EFFECTS
AND SIMPLE MAIN EFFECTS COMPARISONS

Pairs of Principles	Types		
	Conditional	Class	Combined
1 & 2	4.39	3.76	4.08
1 & 3	3.49	3.39	3.44
1 & 4	2.69	2.34	2.51
1 & 5	1.19	.67	.93
2 & 3	.37	.91	.64
2 & 4	1.08	2.08	1.57
2 & 5	2.58	3.72	3.15
3 & 4	.71	1.15	.93
3 & 5	2.21	2.81	2.51
4 & 5	1.50	1.06	1.58

All significant at $p < .05$

TABLE 6
DIFFERENCES FOR NEWMAN-KEULS MAIN EFFECTS COMPARISON
PRINCIPLES AND TYPES

Class						
Conditional	Ordered Means	K ₂	K ₃	K ₄	K ₅	K ₁
		3.45	3.82	4.53	6.03	7.21
K ₂	2.62	- .83*	-1.20*	-1.91*	-3.41*	-4.59*
K ₃	3.53	.08	- .29*	-1.00*	-2.50*	-3.69*
K ₄	4.68	1.23*	.86*	.15	-1.35*	-2.54*
K ₅	6.34	2.89*	2.52*	1.81*	.31*	- .87*
K ₁	7.01	3.56*	3.19*	2.49*	.99*	- .20

$p < .05$

Note: Table shows differences in principle means (Conditional-Class);
negative indicates class has larger mean, hence easier.

TABLE 7
PROPORTIONS OF SUBJECTS MASTERING EACH PRINCIPLE

Principle	Proportion Mastering		
	Conditional	Class	Combined
1	.84	.91	.76
2	.11	.19	.08
3	.21	.33	.17
4	.31	.27	.11
5	.78	.65	.57
All	.00	.00	.00

TABLE 8
PERCENTAGES OF STUDENTS WHO HAD MASTERED EACH OF FIVE PRINCIPLES OF DEDUCTIVE REASONING IN CONDITIONAL FORM (Roberge, 1972)

Grade Level	Principles					n/N	p	Researcher
	1	2	3	4	5			
4	53	2	2	35	28	8/12	.02	Roberge (1970)
5	51	2	3	30	25	5/6	.02	Ennis & Paulus (1965)
4-5	59	0	2	16		5/6	.02	Gardiner (1965)
6	54	0	0	23	28	8/12	.02	Roberge
7	56	3	6	41	45	5/6	.02	Ennis & Paulus
6-7	79	0	2	23		5/6	.02	Gardiner
8	95	0	0	74	81	8/12	.02	Roberge
9	66	4	5	35	40	5/6	.02	Ennis & Paulus
8-9	89	4	4	48		5/6	.02	Gardiner
10	100	19	5	65	82	8/12	.02	Roberge
11	62	3	12	35	58	5/6	.02	Ennis & Paulus
10-11	99	30	20	47		5/6	.02	Gardiner
-----	---	--	--	--	--	---	---	-----
Pre-service Elem. Tchg. Jrs. & Srs.	84	11	21	31	78	6/8	.02	Jansson

n=number of items required for mastery of a principle
N= total number of items for a principle
p= probability of correctly answering at least n items by guessing alone.

ciples of inference. Actually Roberge in this summary lists published results for deductive reasoning only in the conditional form. It is clear both in Roberge’s earlier work (1970) and in the present study that the measured ability to handle a particular principle is different if the argument is presented in the class form. In Table 8 part of Roberge’s (1972, p. 199) table is reproduced, showing the percentages of students who had mastered a given principle at each grade level.

Table 9 presents similar mastery data for deductive reasoning in the class form. The percentage figures in this table come from Ennis and Paulus (1965) and Roberge (1970). Of particular interest here are the variations in findings between researchers, despite the common conceptual framework, on certain principles. Principles 2,3, and 4 all show a wide fluctuation, but it is particularly noteworthy in Principle 3, denying the antecedent. The explanation for such wide fluctuations is not entirely clear, but it must be noted that the results reported for conditional reasoning are much more consistent across researchers. This is a matter that deserves further investigation.

The significant main effects differences in principles and types of reasoning confirm other researchers’ (Ennis & Paulus, 1965; Gardiner, 1965; Paulus, 1967; Roberge, 1970; Roberge & Paulus, 1971) findings. The mastery data found in Table 7 perhaps has the greatest implications for teacher training, particularly if we take seriously other researchers’ suggestions that their findings indicate that implicit instruction in logical skills can begin as early as grade 4. Clearly, large numbers of beginning elementary school teachers are deficient in these rudimentary logical skills of a judgment type.

TABLE 9
PERCENTAGES OF STUDENTS WHO HAD MASTERED EACH OF THE FIVE PRINCIPLES OF DEDUCTIVE REASONING IN CLASS FORM

Grade Level	Principle					n/N	p	Researcher
	1	2	3	4	5			
4	56	5	5	11	34	5/6	.02	Ennis & Paulus (1965)
4	60	2	0	44	35	8/12	.02	Roberge (1970)
6	79	15	15	17	44	5/6	.02	Ennis & Paulus
6	58	4	2	44	28	8/12	.02	Roberge
8	91	21	13	24	55	5/6	.02	Ennis & Paulus
8	82	5	4	79	74	8/12	.02	Roberge
10	92	44	35	37	80	5/6	.02	Ennis & Paulus
10	95	40	21	75	79	8/12	.02	Roberge
12	97	68	47	58	75	5/6	.02	Ennis & Paulus
Pre-service Elem. Tchrs. Jrs. & Srs.	91	19	33	27	65	6/8	.02	Jansson

Efficient instruction in these skills will require further investigation into the instruction of the various principles. Perhaps instruction in one or two of the principles, with appropriate techniques, will be sufficient for mastery of all five. The present study has dealt only with a judgment form, but an extension of the Paulus (1967) study could be made to include assumption finding, first in a situation of logical necessity and later in a more general situation.

The answers to the first three questions of this study have been made in the affirmative, while the results relative to the fourth question reveal the necessity of specific instruction for teachers—maturation alone does not provide the desired skills.

References

- Ennis, R. H. The development of a critical thinking test. Unpublished doctoral dissertation, University of Illinois, 1959.
- Ennis, R. H. A concept of critical thinking. *Harvard Educational Review*, 1962, 32 (Winter), 81-111.
- Ennis, R. H., & Paulus, D. H. *Critical thinking readiness in grades 1-12 (Phase I: Deductive reasoning in adolescence)*. Ithaca, N.Y.: The Cornell Critical Thinking Project, 1965.
- Gardiner, W. L. An investigation of understanding of the meaning of the logical operators in propositional reasoning. Unpublished doctoral dissertation, Cornell University, 1965.
- Hill, S. A. A study of logical abilities in children. Unpublished doctoral dissertation, Stanford University, 1961.
- Howell, E. N. Recognition of selected inference patterns by secondary school mathematics students. Unpublished doctoral dissertation, University of Wisconsin, 1966.
- Jansson, L. C. The development of an instrument to assess critical thinking ability in mathematics. Unpublished doctoral dissertation, Temple University, 1970.
- O'Brien, T. C., & Shapiro, B. J. The development of logical thinking in children. *American Educational Research Journal*, 1968, 5, 531-542.
- O'Brien, T. C., Shapiro, B. J., & Reali, N. C. Logical thinking—Language and context. *Educational Studies in Mathematics* 1971, 4, 201-219.
- Paulus, D. H. A study of children's abilities to deduce and to judge deductions. Unpublished doctoral dissertation, Cornell University, 1967.
- Roberge, J. J. A study of children's abilities to reason with basic principles of deductive reasoning. *American Educational Research Journal*, 1970, 7, 583-596.
- Roberge, J. J. Recent research on the development of children's comprehension of deductive reasoning schemes. *School Science and Mathematics*, 1972, 72, 197-200.
- Roberge, J. J., & Paulus, D. H. Developmental patterns for children's class and conditional reasoning abilities. *Developmental Psychology*, 1971, 4, 191-200.
- Winer, B. J. *Statistical principles in experimental design*. New York: McGraw-Hill, 1962.

MARGARET M. CLIFFORD

University of Iowa

Validity of Expectation—A Developmental Function¹

Validity of self-expectations was found to be an increasing monotonic function of development among first-, third-, and fifth-graders. The relatively high proportion of invalid expectations observed at all grades is attributed to overestimation tendencies and an apparent failure to process task feedback. Given that expectation is a major component of human motivation theories, an investigation of the determinants of valid expectations and means by which they can be manipulated would seem desirable. (Dr. Clifford is Associate Professor in the Division of Educational Psychology, Measurement and Statistics, College of Education, The University of Iowa.)

It has been demonstrated that self-expectation is a predictor of task persistence (Battle, 1965; Feather, 1963) as well as academic achievement (Battle, 1965; Battle, 1966; Biggs & Tinsley, 1970; Keefer, 1969). There is also evidence that expectations affect the tendency to approach tasks (Crandall, Katkovsky & Preston, 1962; Feather, 1966) and solve problems (Feather, 1966; Tyler, 1958). Studies supporting these findings appear to be based on the assumption that individuals generally form relatively accurate expectations. However, there is reason to question the accuracy of children's expectations. They have been found to be affected by adult reactions (Crandall, 1963; Crandall, Good, & Crandall, 1964); it has also been demonstrated that children have a tendency toward wishful, and/or unrealistic expectations (Bialer, 1961; Irwin, 1944; Irwin & Mintzer, 1942; and Marks, 1951). A comparison of related studies by Marks (1951) and Irwin (1953) using similar procedures and measures further suggests that the tendency to overestimate one's ability decreases with an increase in age.

The quality of an expectation can be assessed from at least two perspectives—the conformity of an expectation to subsequent relevant performance and the conformity of an expectation to previous relevant

¹ This research was supported by Public Health Service Grant MH-24091-01. The cooperation of administrators, teachers and pupils from the Clear Creek and Washington Community School Districts in Iowa is gratefully acknowledged.

behaviors. The former concept, here referred to as the Predictive-Value of Expectation, is assumed to be dependent upon the latter concept, subsequently referred to as Validity of Expectation, which is of primary concern in this study.

It is predicted that, among children, Validity of Expectation is an increasing monotonic function of development and that invalid expectations are reflected to a greater extent in overestimation than in underestimation. It is also postulated that accuracy in processing feedback may partially account for Validity of Expectation.

Method

Subjects

The Clear Creek and Washington, Iowa school systems participated in this study. Observations were obtained from four classrooms at each of the first, third, and fifth grade levels. The number of males at each successive level was 47, 42, 49; the number of females was 47, 51, 41. Mean standardized achievement scores from each grade indicated the subjects were representative of average, middle-class students.

Materials

Two types of perceptual judgment problems were designed for this experiment—"string" problems (ST), in which the larger of two rectangular line drawings was to be identified, and "land" problems (LA), in which the larger of two irregularly-shaped areas was to be identified. A spot was located beside each of the members of the various pairs of drawings to be compared. Supposedly, the spot corresponding to a "correct response" was completely erasable, whereas that associated with an "incorrect response" was only partially erasable. In reality, the line lengths or areas were equal, and both spots for any given problem were either erasable or non-erasable.

Three pages, each containing 10 land problems, and another three pages, each containing 10 string problems, were prepared; pages will be identified with numerical subscripts. Four of these six sheets were then assembled in one of two sequences (LA₁, ST₁, LA₂, ST₂ or ST₁, LA₁, ST₂, LA₂). For half of the 4-sheet packets in each sequence the erasable and non-erasable spots were arranged to convey "correct" responses totalling 8, 3, 7, 2, on the four successive pages. The manipulated feedback on the other half of the booklets was 3, 8, 2, 7. In any given booklet, therefore, one type of problem was associated with high feedback while the other was associated with low feedback (*e.g.*, the scores on the land pages were 3 and 2 while the scores on the string pages were 8 and 7). Type of task and level of feedback were counterbalanced. A fifth page, which was either LA₃ or ST₃, completed the booklet. It was used only to elicit a subject's expectation, and was not worked; therefore, no feedback was obtained on this final sheet.

On the back of the booklet, three follow-up items were provided. The subject was asked to indicate (a) whether he found the string or land problems easier, (b) the number of correct responses he made on page one, and (c) whether an example problem, supposedly worked by another student, was "right" or "wrong"; this example depicted a land problem in which an attempted erasure resulted in failure to remove the entire spot (*i.e.*, an incorrect response).

Procedure

Data collection was conducted in intact classrooms to insure a familiar achievement-oriented setting and subjects' typical attentiveness to academic feedback. The compilation of booklets involved a counterbalancing procedure which yielded 16 different booklet forms. A systematic distribution of these forms insured that no two adjacent subjects worked similar pages at the same time. The task was explained with the use of two solvable sample problems, keyed to reflect objectively correct and incorrect responses. Subjects were told the problems in the booklet would be more difficult than the examples, but that they should do their best. Allowing about 1½-2 minutes per page, the experimenter directed the subjects to work the first four pages. An emphasis on a limited time allowance and use of a stopwatch served to economize on time and to minimize student idleness as well as the opportunity for peer comparisons. However, subjects were given time to complete each page.

Upon turning to page 5, subjects were asked to indicate in writing at the top of the page how many of the 10 problems on that page they thought they would answer correctly. Without working the last page, subjects were immediately assisted in reading the follow-up items. An informal debriefing, consisting of answering subjects' questions, was conducted; the experimenter explained that it was not necessary to work the last page, since everyone had worked at least two pages of each type and time was running out.

Expectation Measures

The measures of expectation were the Raw Score Expectation (E_r) recorded on page 5 of the booklet and the Validity of Expectation (E_v) which was a measure of the discrepancy between E_r and the manipulated scores on the two preceding pages of corresponding type (*e.g.*, if page 5 consisted of LA_3 , the mean feedback on LA_1 and LA_2 was compared with the expectation recorded by subject on page 5). The scores obtained on the pages similar in type to page 5 are subsequently referred to as "relevant feedback."

Validity of Expectation was coded as follows: A score of 1 was assigned to a subject whose Raw Score Expectation was two or more points lower than his lower relevant feedback (*i.e.*, E_r of 5 or less for subjects with relevant feedback of 7 and 8, and 0 for subjects with relevant feedback of 2 and 3) or 2 or more points higher than the higher relevant feedback (*i.e.*, E_r of 10 for subjects with relevant feedback of 7 and 8, and E_r of 5 or more for subjects with relevant feedback of 2 and 3). A score of 2 was assigned to a subject whose expectation was one point lower than his lower relevant feedback or one point higher than his higher relevant feedback. A score of 3 was assigned to a subject whose expectation coincided with either of his two relevant feedback scores (*i.e.*, 7 or 8 in the high and 2 or 3 in the low feedback condition).

Results²

A Sex x Grade analysis of variance of Validity of Expectation yielded a Grade effect, $F(2, 271) = 4.74$; $p < .01$. Means indicated an increase in Validity of Expectation with grade; they were 1.34, 1.41, 1.67 for first, third, and fifth respectively, and thus offered support for the prediction that there is an increasing monotonic relationship between these variables. However, Scheffé tests between adjacent grades were not significant at the .05 level. Neither the Sex effect nor the Grade x Sex interaction was significant.

A Sex x Grade analysis of variance of Raw Score Expectation also resulted in a Grade effect, $F(2, 271) = 10.77$; $p < .001$, in the absence of a Sex or interaction effect. The nature of this effect is, as expected, the reverse of that found for Validity of Expectation; means were 7.78, 7.04, and 6.30 for first, third and fifth respectively. Whereas Validity of Expectation increases with Grade, Raw Score Expectation decreases. Given that the average relevant feedback across subjects in any grade was 5, and the mean Raw Score Expectation exceeds 6 at every grade level, these data offer support for the second hypothesis that invalid expectations are reflected more in overestimations than in underestimations.

Once again, Scheffé tests between adjacent grades failed to indicate significant differences. However, the frequency with which subjects stated an expected score of 10 on page 5 supported the predicted developmental trend. This maximum score (unobtained by any subject on previous trials) was recorded by 44% of the subjects in the first grade, 20% in the third grade, and 3% in the fifth grade. The procedure for testing proportions reported in Blommers and Lindquist (1960) indicated all pairwise comparisons between these percentages were significant beyond the .001 level.

Additional analyses of variance were performed to investigate whether these data were dependent upon subjects stating an expectancy on a task associated with "success" vs "failure"; half the subjects predicted for a task on which they had received scores of 2 and 3 (Low Prediction Page), whereas half predicted for a task on which they had previously scored 7 and 8 (High Prediction Page). Two 2 x 3 (Prediction Page by Grade) analyses of variance showed that aside from the Grade effect previously reported on both Validity of Expectation and Raw Score Expectation measures, only one significant effect resulted; namely, Validity of Expectation was higher in the High as opposed to the Low Prediction Page condition. Given the pervading tendency to overestimate, this latter finding may reflect nothing more than a ceiling effect; the discrepancy between the maximum possible score and High relevant feedback cannot be as great as the discrepancy between the maximum possible score and Low relevant feedback.

A more detailed assessment of subjects' expectations under these two conditions is presented in Table 1. It suggests that High expectations of fifth-graders, unlike those of first- and third-graders, were associated almost exclusively with failure-oriented tasks. Using the procedure reported in Blommers and Lindquist (1960), the proportion of fifth-graders who recorded expectations of 9 or 10 was significantly greater ($p < .002$) for the

² The univariate analyses reported represent a least squares solution which allows for disproportionate cell frequencies.

TABLE 1
PERCENTAGE OF LOW, ACCURATE, AND HIGH EXPECTATIONS*
GIVEN BY THE HIGH AND LOW PREDICTION PAGE SUBJECTS

Prediction Page Type	Grade	N	Expectation		
			Low	Accurate	High
High	1	48	31	25	44
Low	1	46	04	00	96
High	3	45	44	22	33
Low	3	48	00	04	96
High	5	46	50	43	07
Low	5	44	00	05	95

* Low: below low relevant feedback
Accurate: equal to relevant feedback
High: above high relevant feedback

| pairwise test, $p < .05$ || pairwise test, $p < .01$

Low as opposed to the High Prediction Page. Comparable tests of proportions for first- and third-graders showed nonsignificant differences.

Table 1 also indicates that accurate expectations occurred almost exclusively in the High as opposed to the Low Prediction Page. Furthermore, whereas inaccurate responses in the High Prediction Page condition were rather evenly distributed between Low and High expectations for first- and third-graders there was a noticeable tendency for fifth-graders to express low more frequently than high expectations in this condition, 50% *vs* 7%, $p < .01$. (Several obviously significant contrasts, assumed to be a phenomena of overestimation or ceiling effects, which were discussed earlier have not been indicated in this table.)

A summary of the accuracy with which subjects responded to the three follow-up items is presented in Table 2. Whereas the first and third items allowed for only one of two responses (correct or incorrect), the second item allowed for one of 10 responses. A response to this item was considered accurate if the number of correct responses reported by the subject was no

TABLE 2
PERCENT OF SUBJECTS WITH ACCURATE RESPONSES ON FOLLOW-UP ITEMS

Items	Grade		
	1	3	5
Task Difficulty	52	67	80
Correct Responses	65	86	89
Judgment of Example	82	81	89

| pairwise test, $p < .05$ || pairwise test, $p < .01$

more than one point discrepant from the manipulated feedback on his first page. (This one-point discrepancy margin was in acknowledgment of the possibility of inadvertent errors in counting.) Any other response to this item was classified as inaccurate.

The increase in accuracy with increase in grade level found for item 1 lends credence to the speculation that accuracy in processing feedback may be one explanation for the developmental function of Validity of Expectation. The accuracy pattern for item 2 (assumed to require less complex skills than item 1) suggests that accuracy develops more rapidly with age, the less complex the feedback-processing task. The relatively high accuracy on item 3 and the consistency across developmental levels suggests most subjects could, indeed, discriminate between correct and incorrect feedback. Whereas third- and fifth-graders showed no difference in accuracy regardless of whether they interpreted their responses (item 2) or another's response (item 3), first-graders showed significantly greater accuracy in interpreting the feedback on the example as compared to their own first page. Furthermore, 85% of the first-graders who responded inaccurately on item two reported an inflated score for themselves.

Discussion

The developmental function demonstrated in this experiment suggests that if, as frequently assumed, valid expectations are a major component of human motivation, the manipulation of Validity of Expectation in children may warrant attention. While age emerges as the most obvious factor accounting for variability of Validity of Expectation, it obviously precludes manipulation. If, however, one postulates that this developmental function of Validity of Expectation can alternatively be explained by an increasing monotonic relationship between Age and accuracy in processing feedback, training on skills such as perception, recognition, discrimination, association, and integration would appear to be viable techniques by which to affect Validity of Expectation. More specifically, it is postulated that providing concrete, visual feedback (e.g., charts, tallies) on performance, will increase Validity of Expectation. Similarly self-scoring and recording of performance may also enhance Validity of Expectation for children.

Furthermore, if one postulates that this developmental function of Validity of Expectation is partially explained by a decreasing monotonic function between age and an overestimation tendency, reality training, (e.g., emphasizing the difference between what one would like to have happen and what one thinks will happen in a given situation) may provide a means of manipulating Validity of Expectation. As suggested in the results, the responses on the follow-up items seem to lend credence to these postulations. Finally, given that even for fifth graders Validity of Expectation, as coded in this study, does not exceed 1.8 on a 3-point scale, it seems advisable to reinvestigate this variable on an extended developmental continuum.

Validity of Expectation—A Developmental Function

References

- Battle, E. Motivational determinants of academic task persistence. *Journal of Personality and Social Psychology*, 1965, 2, 209-218.
- Battle, E. Motivational determinants of academic competence. *Journal of Personality and Social Psychology*, 1966, 4, 634-642.
- Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. *Journal of Personality*, 1961, 29, 303-320.
- Biggs, D. A., & Tinsley, D. J. Student-made academic predictions. *Journal of Educational Research*, 1970, 63, 195-197.
- Blommers, P., & Lindquist, E. F. *Elementary statistical methods*. Houghton-Mifflin Company, Boston, 1960.
- Crandall, V. C. Reinforcement effects of adult reactions and nonreactions on children's achievement expectations. *Child Development*, 1963, 34, 335-354.
- Crandall, V. C., Good, S., & Crandall, V. J. Reinforcement effects of adult reactions and nonreactions on children's achievement expectations: A replication study. *Child Development*, 1964, 35, 485-497.
- Crandall, V. J., Katkovsky, W., & Preston, A. Motivational and ability determinants of young children's intellectual achievement behaviors. *Child Development*, 1962, 33, 643-661.
- Feather, N. T. Persistence at a difficult task with alternate task of intermediate difficulty. *Journal of Abnormal and Social Psychology*, 1963, 66, 604-609.
- Feather, N. T. Effects of prior success and failure on expectations of success and subsequent performance. *Journal of Personality and Social Psychology*, 1966, 3, 287-298.
- Irwin, F. W., & Mintzer, M. G. Effect of differences in instructions and motivation upon measures of the level of aspiration. *American Journal of Psychology*, 1942, 55, 400-416.
- Irwin, F. W. The realism of expectations. *Psychological Review*, 1944, 51, 120-126.
- Irwin, F. W. Stated expectations as functions of probability and desirability of outcomes. *Journal of Personality*, 1953, 21, 329-335.
- Keefer, K. E. Self-prediction of academic achievement by college students. *Journal of Educational Research*, 1969, 63, 53-56.
- Marks, R. W. The effect of probability, desirability, and privilege on the stated expectations of children. *Journal of Personality*, 1951, 19, 332-351.
- Tyler, B. B. Expectancy for eventual success as a factor in problem solving behavior. *Journal of Educational Psychology*, 1958, 49, 166-172.

LOIS FOSTER

and

MARY NIXON

University of Alberta

The Interview Reassessed

In this paper the interview is discussed as problematic. This approach leads to an examination of the interview as "a product of social arrangements" (Young, 1972, p. 194) rather than a prescription for strategies of operation. Such discussion may raise new questions with the possibility of new answers. (Lois Foster and Mary Nixon are graduate students in the Faculty of Education, The University of Alberta.)

The recent trend in sociology of education which has attacked the positivistic stance of much of the reported research has been accompanied by a renewal of interest in other approaches (Young, 1971; Hopper, 1971; Filmer *et al.*, 1972). As a result there has been an increasing emphasis on research based on the interpretative paradigm which challenges the dominance of highly statistical research oriented to the normative paradigm (Colfax & Roach, 1971; Brown, 1973). The dominant theory underlying the normative paradigm is structural functionalism which seeks an understanding of a whole system by analysing the function served by each of the individual parts. The interpretative paradigm derives its orientation from such theoretical bases as symbolic interactionism and phenomenology which focus upon the interactions of people. For the researcher, working within the normative perspective, objectivity as inter-subjective testability is a major goal. In contrast, the interpretative paradigm takes account of the interaction between researcher and researched and incorporates it within the research design.

At the present time there is a tendency to look upon these two paradigms as discrete. Yet a careful appraisal of the positivistic work done by such sociologists as Halsey, Floud and Anderson (1961) reveals an appreciation for the underlying questions raised by the interpretative approach (Eggleston, 1973). This would suggest the possibility, at least, that these two approaches need not be mutually exclusive.

The interview is probably the most adaptive instrument of data collec-

tion. The researcher may place the emphasis deliberately upon the interaction between researcher and participants or the emphasis may be placed on the elimination of reactive effects by the use of *a priori* hypotheses and closed response categories. An alternative to this dichotomy is a format which combines predetermined questions designed to elicit specific information with spontaneous questions that arise from the interaction between the researcher and the participant. Surprisingly, the interview method, although it is as much the result of research orientation as the most structured, written survey technique, does not appear to have come in for as much criticism as the questionnaire method.

Opportunities for some experience in diverse interview situations have suggested the need to examine the interview as problematic. To define the interview as problematic is to define it as "a product of social arrangements," (Young, 1972, p. 194) that is, as a result of the activities of people. Textbooks such as Selltiz *et al.* (1965), Phillips (1966), Good (1966), treat the interview as a conventional typification comprised of "taken for granted" definitions of phenomena which help to make the everyday world meaningful to the actor (Schutz, 1962). The assumption of a common definition or a common research purpose may not be warranted. Methodology courses are frequently little better with their suggestion that the interview is the "human" replacement of the written questionnaire. The implied criterion for the use of the interview is often the extent of the funds available to the researcher. Those studies which have substantial funding may choose to use the interview, but as an addition to an extensive printed survey. Modestly funded studies may select the interview, particularly the telephone interview, as an inexpensive method of data collection. Another criterion for the use of the interview is to improve an unacceptably low response rate obtained by a written questionnaire.

Textbook writers tend toward an atheoretical treatment of the interview. Indeed it is often dismissed. For example, one recent text on survey and research methods devotes fourteen out of a total of three hundred and seventy-two pages to general rules for interviewing (Babbie, 1973, pp. 171-185). Such information as "the interviewer should be relaxed and friendly without being too clinging" (Babbie, 1973, p. 173), that he should be familiar with the questions (Babbie, 1973, p. 174), or that the time of day should be carefully controlled (Travers, 1969, p. 236), are practical hints of a mundane nature. Less helpful is the explanation that "the term 'interview' is constructed out of two words, 'inter' and 'view'" (Lofland, 1971, p. 75), while lack of solutions detracts from the utility of Glazer's warnings of the problems to be encountered in interview research (Glazer, 1972, p. 25). Those who seek a theoretical treatment of the interview are likely to be disappointed.

It is not always recognized that the structured interview differs only in format from the structured questionnaire. In this type of research, meaning is imposed on the situation by the questions asked and by the researcher's attitude which filters out information that is seen as peripheral to the phenomenon under investigation. The problematic approach would raise questions of (a) the influence of value and belief systems, (b) ways of defining legitimate data, and (c) the emphasis and direction of the research report.

The adoption of a problematic approach stimulates a rethinking of the interview, the justification for its use and the legitimacy of the data outcomes. Interview research ranging from children without choice as to their participation to influentials who volunteered their resources of time and knowledge provided the experiential base. Yet this diversity lent itself to a classification on two dimensions: influential . . . non-influential and selected . . . self-selected. The definitions of influential . . . non-influential and selected . . . self-selected are not mutually exclusive categories since an individual who can be termed "self-selected" in one research situation may be constrained to participate in another, thus becoming, in this sense "selected." In general, individuals or groups may choose whether they respond to the researcher's request that they participate in a study but for certain groups (*e.g.* children) and in certain situations (*e.g.* government sponsored studies) the decision to participate is not voluntary. Societal status has an important bearing not only on an individual's perception of his position relative to the implications of the study but also on the researcher's perceptions which affect the design of the study.

Recognition of the importance of values, those of the researcher and the researched, both influentials and non-influentials, transforms assumptions into questions. The assumption that reactive effects are necessarily sources of error appears in such authoritative texts as that of Rundle and McGrath (1972, p. 184). A design attempting to reduce these "sources of error" will tend toward a format which emphasizes prescribed wording. A positive view of reactive effects would contest this assumption. It raises questions as to qualitative ways of knowing. The researcher faces a decision as to whether he will attempt a more complex design that will incorporate qualitative data or whether he will restrict himself to quantifiable data which can be manipulated with appropriate statistical tools.

Even if the researcher is content with quantifiable data he cannot with impunity ignore the participants' value systems since values and beliefs influence the responses. It would indeed be a dangerous assumption that there can exist a question that, for all individuals and groups, could be entirely value free. A structured format, however, minimizes the intrusion of values and enhances a study's replicability. The interpretative approach to research design takes into account qualitative data but its objectivity may be more readily challenged.

The researcher, influenced by the positivistic approach, may too readily make the assumptions: (1) that a sophisticated statistical treatment of his data leads to a sophisticated understanding of his topic, and (2) that questions will elicit answers reflecting a participant's "true" position on the concern of the study, that is, that responses that are reliable over time are indeed valid. Those not committed to a normative approach would suggest: (1) that a qualitative treatment of data is a valuable contribution to understanding, and (2) that the participants' perceptions become the concern of the study. In practice the first approach will tend toward an interview situation where the participant is required to give answers structured to allow categorization of responses suitable for transfer to computer cards. The second approach is more open-ended and thus the data that is gathered frequently eludes a ready classification. Constraints are imposed on the methodology by the researcher's perception of the influential or non-

influential status of the interviewees. Legitimate data takes a variety of forms. In certain cases, for example, the collection of biographical data, the use of the second approach would be time consuming and probably unnecessary.

The reports which result from these two perspectives will differ in emphasis and direction but both are subject to bias. The emphasis of normative research is on *a priori* hypotheses which are to be supported or rejected by statistical data analysis. This imposes defined limits on the research report because once the findings have been presented the conclusions are almost self-evident. However, shifting cutting points and levels of significance manipulate the findings in the direction of the researcher's bias. The interpretative approach makes no formal statement of hypotheses in advance and therefore the analysis of the data presented in the report is subject to greater variation. Recognizing that bias cannot be eliminated, the researcher attempts to analyze his own position in relation to the data. Thus intuitive rather than quantifiable ways of knowing take precedence in this approach.

Reassessment of the interview arose from the variety of definitions derived from diverse interview situations. Considering the interview as problematic permits a scrutiny of assumptions and strategies which are generally left unquestioned. Yet, as of the present, the interpretative approach lacks operational definitions of concepts. Interviews conducted from this perspective would be different from, but not necessarily better than, those conducted within the normative framework. With more examples of the use of the interpretative approach, the result may be that the two perspectives are profitably incorporated into a new research paradigm.

References

- Babbie, E. *Survey research methods*. California: Wadsworth Publishing Co., Inc., 1973.
- Brown, R. (Ed.) *Knowledge, education and culture*. London: Tavistock Publishing Co., 1973.
- Colfax, T., & Roach, J. (Eds.) *Radical sociology*. New York: Basic Books, 1971.
- Eggleston, S. *Sociology and the sociology of education*. University of Alberta, November, 1973. (Mimeo.).
- Filmer, P., Phillipson, M., Silverman, D., & Walsh, B. *New directions in sociological theory*. London: Collier-Macmillan Publishing Co., 1972.
- Halsey, A., Floud, J., & Anderson, C. *Education, economy and society*. Glencoe, Ill.: The Free Press, 1961.
- Glazer, M. *The research adventure*. New York: Random House, 1972.
- Good, C. V. *Essentials of educational research, methodology and design*. New York: Appleton-Century Crofts, 1966.
- Hopper, E. (Ed.) *Readings in the theory of educational systems*. London: Hutchinson, 1971.
- Lofland, J. *Analyzing social settings*. California: Belmont Publishing Company, 1971.
- Phillips, B. *Social research: Strategy and tactics*. New York: The Macmillan Co., 1966.
- Rundle, P., & McGrath, J. *Research on human behaviour*. New York: Holt, Rinehart and Winston Publishing Company, 1972.

- Schutz, A. *Collected papers Volume 1: The problem of social reality*. M. Natanson (Ed.) The Hague: Martinus Nijhoff, 1962.
- Selltiz, C., Jahoda, M., Deutsch, M., & Cook, S. *Research methods in social relations*. New York: Holt, Rinehart and Winston, 1964.
- Travers, R. *An introduction to educational research*. New York: The Macmillan Company, 1969.
- Young, M. F. D. *Knowledge and control*. London: Collier-Macmillan Publishing Company, 1971.

V. FROESE

The University of Manitoba

Does The Dolch Do?

Considerable controversy has been generated recently with regards to the usefulness of various word lists. The present study deals with students' actual recognition of words on the Dolch 220 list and another comparable list drawn from the Kučera/Francis corpus.

The Dolch 220 appears to differentiate between grades and IQ levels and correlated more highly with a criterion measure administered five months later. Both tests did not differentiate between the sexes or on socioeconomic status. (V. Froese is Assistant Professor in the Department of Curriculum, Faculty of Education, The University of Manitoba.)

Considerable interest has been generated recently in word lists by Johnson (1971, pp. 449, 451), Johns (1971), Otto and Chester (1972), Durr (1973), Harris and Jacobson (1973-74) and Pennock (1974). Consequently, the inveterate Dolch 220 has come under close scrutiny. Although the pervasiveness of this list is not precisely known, Jacobs (1967, p. 2) concluded from his interview of teachers that (1) the Dolch lists are familiar to most teachers in remedial and primary reading, (2) the use of the lists may be greatest in the area of remedial reading, and (3) few people are aware of the origins and ages of the lists.

The Dolch 220 also known as *The Basic Sight Word Test*, was based on vocabulary studies of the 1920's and is still used in its original form. The current form was copyrighted by Garrard Publishing Company in 1942.

It seems logical to question the usefulness of this list simply because of its age and it is not surprising to find comments such as the following:

The Dolch List, as a corpus, has outlived its usefulness and . . . a more adequate substitute is available (Johnson, 1971, pp. 449, 451).

Yet the fact is that in addition to becoming somewhat passé over more than three decades, the list [Dolch 220] had a questionable empirical base in the first place (Otto & Chester, 1972, p. 435).

This paper describes a study which was designed to investigate the current usefulness of the Dolch 220. It was decided to test second and third grade children's actual recognition of the Dolch words and compare the

results to scores on a test suggested by Johnson (1971). It is important to note, then, that this study deals with *recognition* of words and is not a study based on text demand (frequency counts, occurrence in basal readers, etc.). To provide some indication of the test's predictive validity, the *Botel Word Opposites Test* (1966) was administered five months later. This test was used since it correlated highly, from 0.82 to 0.90, with other comprehension measures (Botel, 1970).

Before describing the results it is instructive to consider the research on word recognition tests. It is essential to keep in mind four factors influencing vocabulary study: (1) the source—whether from children or adults, (2) the modality—whether oral or written, (3) the method of collection—whether frequency counts, free-association, from basal readers, tested knowledge, adult judgment, or some combination, and (4) the period or time of collection—whether in the 1920's or late 1960's and early 1970's, etc.

Related Research

The importance of a basic sight vocabulary was clearly articulated by Dolch. His list contained the words which indicated the structural relationships between members of the form classes (nouns, verbs, adjectives, adverbs) as well as a number of "irregular" words, both being essential to all subject matter material. Furthermore, Dolch (1960) demonstrated that the basic sight vocabulary comprised over one-half of the running words in the subjects of Reading, Arithmetic, Geography, and History.

Johns (1971) replicated this analysis for five current reading series and found that the Dolch words still comprised over one-half of the running words, although the percentage at each grade level (1-6) was somewhat lower. Johnson and Barrett (1972) reported similar findings for Johnson's list which contains many of the same words.

The similarity of these counts appears to indicate that publishers take into account the various studies of basic vocabularies or that they are governed by the fact that these words are necessary to all writing.

Vocabulary research has also influenced the measurement of reading difficulty, since according to Klare (1963), vocabulary accounts for the greatest amount of variance in readability formulas. And since readability formulas are also utilized to control basal reader material, the cycle of interdependent factors pointed out by Froese (1971) is completed.

A further problem related to the study of vocabulary is an environmental one—changes in technology, advancements in the sciences, and cultural influences have their concomitant language influences. Jacobs (1967) has documented one aspect of this change by replicating the Buckingham-Dolch Free-Association Word Study. He concluded that the 1926 Free-Association List is no longer representative of current students' vocabularies in terms of list content, grade-level assignment, and student performance.

Johnson (1971, p. 451) also referred to this change when he stated: "Of the 220 words on the Dolch Primary Word List, representative of the 1920's, *eighty-two* words or *37 percent* are not among the most frequently occurring 220 words in the Kučera and Francis corpus compiled in the 1960's." This statement must, however, be interpreted carefully since the Kučera and

Francis study is based entirely on frequency of occurrence, includes nouns, and is derived from adult material—all of these are facets not common to the Dolch 220.

Lexicography also is concerned with language shifts and regular revisions of dictionaries are practically assumed. The motivation for the *Word Frequency Book* (1971) rests on these grounds but breaks new ground in applying computerized techniques to the vocabulary of children in grades 3 through 9 (and was likely inspired by Kučera's work with adult vocabulary).

A British study by Edwards and Gibbon (1964) examined the written vocabulary of children ages six to eight years and found that it expressed "exciting scientific wonders," that their lists overlapped but were not closely alike the American lists (of Rinsland, Thorndike, Gates and Dolch), and that the earlier lists were "out-of-date in respect of the scientific advances of recent years."

A further consideration to be taken into account when studying vocabulary is the shortcoming of frequency counts alone. Familiarity, meaning, and comprehension do not bear a one-to-one relationship to frequency as pointed out by Klare (1963). Dolch (1951) noted this and suggested that the factors of opportunity of use and emotional set are the causes of the discrepancy between frequency of use and tested word knowledge.

The Study

The purpose of this study was to compare students' tested recognition of the Dolch 220 words with their responses to the 200 highest frequency words found by Kučera and Francis in their *Computational Analysis of Present-Day American English* (1967). The scores on the Dolch test were also compared to current student I.Q.'s and comprehension scores as tested five months later by the *Botel Word Opposites Test*.

Subjects for the study were 155 second graders and 179 third graders from classrooms in four schools in a large suburban school division in Winnipeg, Manitoba, Canada. Winnipeg is a city with a population of approximately 560,000. The factors of grade level, I.Q. level, socioeconomic level, and sex were considered.

The findings are presented in Table 1.

From a statistical analysis, it was found that significant differences occurred at the .05 level between tests, between grades, and among I.Q. levels on the Dolch test. On the Kučera/Francis 220 only the difference between middle and lower I.Q.'s was significant. The differences between socioeconomic levels and sexes did not reach significance.

A further analysis consisted of correlating the I.Q.'s, *Botel Word Opposites* scores, and the scores on the two vocabulary test. The results are presented in Table 2.

All correlation coefficients except those marked NS (non-significant) are significant at the .01 level.

The fact that few grade three students made errors on these tests obviously resulted in a ceiling effect and hence the non-significant findings.

However, the correlations between the Dolch 220 and I.Q., as well as

TABLE 1
MEANS AND STANDARD DEVIATIONS OF SELECTED VARIABLES
ON THE DOLCH 220 and the KUČERA/FRANCIS 220
HIGHEST FREQUENCY WORDS

Variable	Subjects	Means and Standard Deviations	
		Dolch 220	Kučera/Francis 220
Grade			
2	155	205 (19.5)	211 (17.7)
3	179	217 (5.1)	218 (3.8)
I.Q.			
Upper	118	216 (7.6)	218 (5.5)
Middle	151	212 (14.6)	216 (11.9)
Lower	66	203 (21.2)	209 (20.4)
SES			
Middle	181	213 (12.2)	216 (8.3)
Lower	154	210 (17.7)	214 (16.7)
Sex			
Boys	164	210 (16.0)	215 (11.6)
Girls	170	213 (14.0)	216 (14.2)

TABLE 2
INTERCORRELATION FOR THE DOLCH 220, THE KUČERA/FRANCIS 220,
THE BOTEL WORD OPPOSITES TEST AND OTIS I.Q.'S

	Dolch 220	Kučera/Francis 220
Grade 2		
I.Q.	.38	.30
Botel	.51	.42
Kučera/Francis	.84	-
Grade 3		
I.Q.	.09 NS	.01 NS
Botel	.38	.20 NS
Kučera/Francis	.82	-

between the Dolch 220 and the delayed Botel test are considerably higher than the same comparison for the Kučera/Francis 220.

Conclusions

A consideration of the data leads to the following conclusions about the Dolch 220:

1. The Dolch 220 test appears to differentiate between the grade levels considered.
2. The Dolch 220 test appears to differentiate between all three I.Q. levels considered.
3. The Dolch 220 test correlates higher with I.Q. scores than the other test considered.

4. The Dolch 220 test correlates higher with a comprehension measure administered five months later than does the other test considered.

In consideration of the original question, the conclusion appears to be that the Dolch does do!

This study was supported by the University of Manitoba Research Board.

References

- The Basic Sight Word Test*, Dolch, E. W. Champaign, Ill.: Garrard Publishing, 1942.
- Botel, M. *The Botel reading inventory*. Chicago: Follett Educational Corporation, 1966.
- Botel, M., Bradley, J., & Kashuba, M. The validity of informal reading testing. In W. K. Durr (Ed.), *Reading difficulties: Diagnosis, correction, and remediation*. Newark, Delaware: IRA, 1970, 85-103.
- Carroll, J. B., Davies, P., & Richman, B. (Eds.) *Word frequency book*. New York: Houghton-Mifflin and American Heritage, 1971.
- Dolch, E. W. Tested word knowledge vs. frequency counts. *Journal of Educational Research*, 1951, 44 (February), 457-470.
- Dolch, E. W. *Teaching primary reading*. (3rd ed.) Champaign, Ill.: Garrard Press, 1960, 258.
- The Dolch list re-examined—a further comment. In Crossfire, *The Reading Teacher*, 1972, 25 (March), 533.
- Durr, W. K. Computer study of high frequency words in popular trade juveniles. *The Reading Teacher*, 1973, 27 (October), 37-42.
- Edwards, R. P. A., & Gibbon, V. *Words children use: A survey of the words used by children in infants' schools with the resultant graded vocabulary*. London: Burke Publishing, 1964, 105.
- Froese, V. Word recognition tests: Are they useful beyond grade three? *The Reading Teacher*, 1971, 24 (February), 434.
- Harris, J. A., & Jacobson, M. D. Some comparisons between basic elementary reading vocabularies and other word lists. *Reading Research Quarterly*, 1973-1974, 9, 87-109.
- Jacobs, H. D. A free-association word list for the Willamette Valley. Bulletin of the Oregon School Study Council, Vol. II, University of Oregon, 1967.
- Johns, J. L. The Dolch Basic Word List—Then and now. *Journal of Reading Behavior*, 1971, 3 (Fall), 35-40.
- Johnson, D. D. The Dolch list re-examined. *The Reading Teacher*, 1971, 24 (February), 449, 451.
- Johnson, D. D., & Barrett, T. C. Johnson's basic vocabulary for beginning reading and current basal readers: Are they compatible? *Journal of Reading Behavior*, 1972, 4 (Fall), 1-11.
- Klare, G. D. *The measurement of readability*. Ames, Iowa: Iowa State University Press, 1963, 164-169.
- Kučera, H., & Francis, W. N. *Computational analysis of present-day American English*. Providence, Rhode Island: Brown University Press, 1967.
- Otto, W., & Chester, R. Sight words for beginning readers. *The Journal of Educational Research*, 1972, 65 (July-August), 435-443.
- Pennock, C. D. Quick word lists for Canadian readers. *The Alberta Journal of Educational Research*, 1974, 20 (1), 8-13.

DOYAL NELSON

and

DAIYO SAWADA

University of Alberta

Studying Problem Solving Behavior in Young Children—Some Methodological Considerations

This is a report of certain aspects of a research project, "The Nature and Development of Problem Solving Behavior in Early Childhood." In the project, all raw data were recorded on video tape. Some considerations to be kept in mind when video tape is used in research are discussed. Some preliminary but tentative outcomes of the study are also reported. (Dr. Nelson is Professor, and Dr. Sawada is Associate Professor in the Department of Elementary Education, The University of Alberta.)

Mathematics instruction at the early childhood level should be aimed at helping children solve problems associated with situations and events they find themselves in from day-to-day. This problem solving cannot be the usual kind in which the solver seeks to find some mathematical expression to fit reality. The symbol system of the child is too incomplete to process problems of any consequence in this way. However, it should be possible to get children involved in solving problems which require the use of a variety of mathematical processes and which require the use of no symbols at all.

One difficulty with the assumption just stated is that not much is known about creating such problems for children and not much is known about the subsequent behavior of children in solving them. A research study is now under way at the University of Alberta, Department of Elementary Education, which addresses itself to these and other aspects of this matter.¹

The purpose of this paper is:

1. To discuss the guidelines used in the study to develop "good" problems for young children and to describe the nature of some problems so developed.

¹ This research was supported by a Canada Council research grant; Director: Doyal Nelson; Co-investigator: Daiyo Sawada.

2. To discuss the video tape recorder as a device for gathering data on the problem solving behavior of young children.

The Guidelines

Some years ago, a colleague, Dr. Joan Kirkpatrick, and the investigators set about to construct a model which could be used to create "good" problems for young children. We wanted to be reasonably sure that if the criteria stated in the model were adhered to, that the resulting problem or problem situation would indeed stimulate problem solving activity on the part of the child. The criteria which finally emerged as follows:

Criteria for Creating Good Problems

- A. *The problem should be of significance mathematically.* It is for the potential of the situation as a vehicle for the development of mathematical ideas that we choose a particular problem situation or family of situations above all others.
- B. *The situation in which the problem occurs should involve real objects or obvious simulations of real objects.* The main consideration here is that it be comprehensible to the child and easily related to his world of reality.
- C. *The problem situation should capture the interest of the child either because of the nature of the materials, the situation itself, the transformations the child can impose on the materials, or because of some combination of these factors.*
- D. *The problem should require the child to make moves or transformations or modifications with or in the materials.* It is difficult to overemphasize the role of action in early childhood learning. Most of the mathematical models we are interested in at this level are what might be called "action models."
- E. *Wherever possible, problems should be chosen which offer opportunities for different levels of solution.* If the child can move immediately from the problem situation to an expression of its mathematical structure it is not a problem.
- F. *Whatever situation is chosen as the particular vehicle for the problems, it should be possible to create other situations which have the same mathematical structure.* Another way of saying this is that the same problem should have many physical embodiments. It may not be possible for a child to generalize a solution to a certain kind of problem until the problem has come up in a variety of situations.
- G. *Finally, the child should be convinced he can solve the problem and should know when he has a solution for it.* If the child is somehow required to react with or transform materials used in problem situations it is usually easy to determine whether the problem meets the criterion or not.

The model is still a crude one but its application has been reasonably effective in providing direction for the creation of good problem situations for young children. Refinements are being made to the model as it is being used. The following section provides some samples of problems and problem situations which have resulted from the application of the

guidelines in the model. These were developed by a group of graduate students and staff over a short period of time.

Some Problem Situations

In this section, six problem situations are described as examples of the applications of the model criteria.

1. *Fold Out Shapes*. One important thing for children to learn is how three dimensional things can be represented in two dimensions. It is possible, of course, to disassemble and reassemble boxes of various shapes or to look at projections of these shapes or to trace around their faces. None of these activities, however, gives the child the control that our criteria suggest.

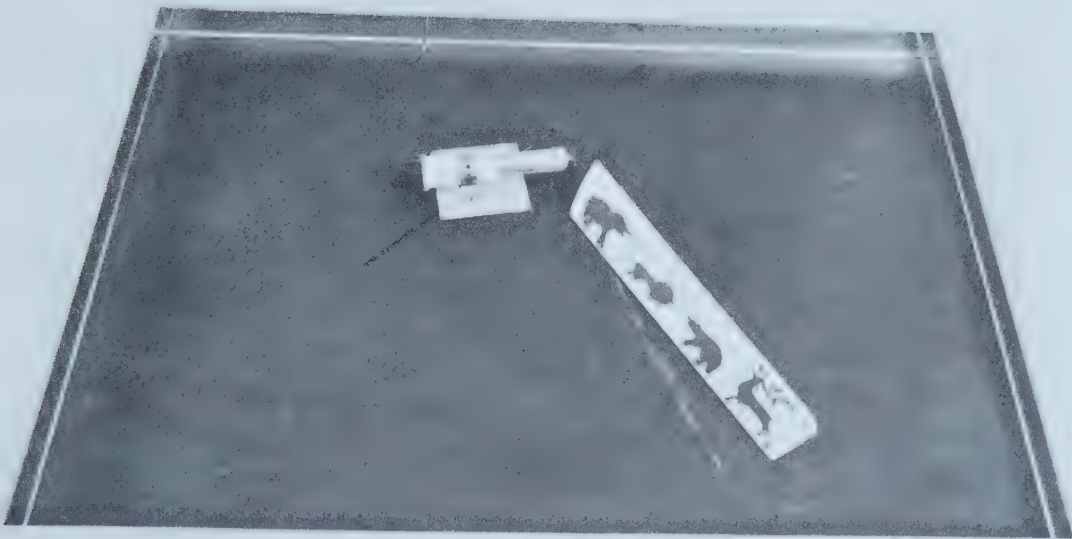
To overcome this difficulty, squares, rectangles, triangles, pentagons, and hexagons all with dimensions of ten centimeters or some multiple of ten centimeters were cut from large sheets of plexiglas. Then, on alternate edges, the positive (hooked) side of Velcro was glued. The negative (plush) side was glued to the remaining edges. (Of course, two positive or two negative faces would show up on one pair of adjacent edges in the triangles and pentagons.) Now the child could be provided with a set of faces which would make a solid join when put together but at the same time would be easy to take apart. A set of the Fold Out Shapes is shown in Photograph 1.



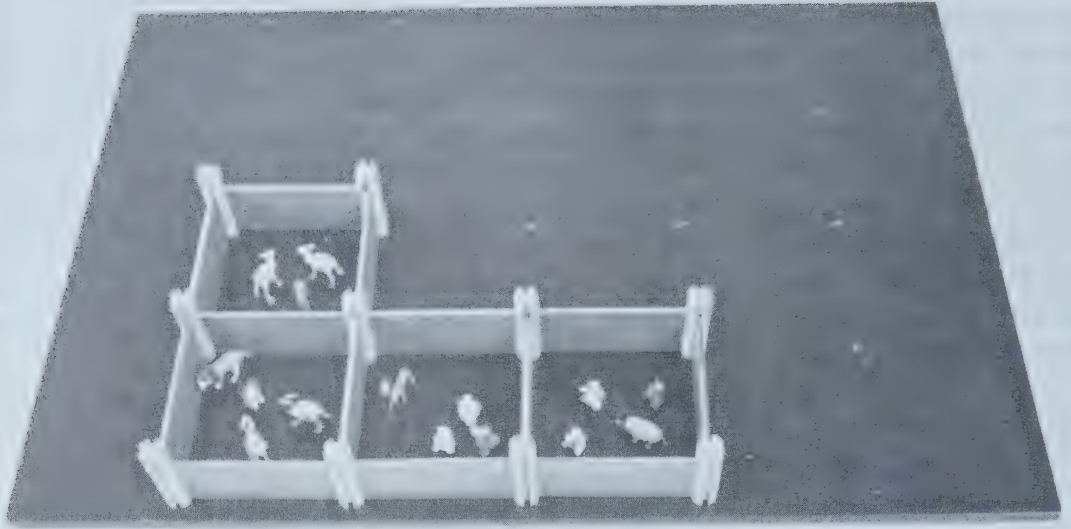
The kind of problems this material opens up for the child is very extensive. With sufficient pieces the Platonic solids can be constructed as well as a number of interesting, compound shapes.

2. *Mirror Reflection.* The next photograph shows the layout used to help children trace out the nature of a reflected beam of light. The board itself has shallow sides to which are attached mirrors. A point source of light was mounted on a carrier which could be moved along a groove and rotated 360°. There was a movable target which consisted of four animals. The basic problem was for the child to reflect the light off one mirror so that it fell on a designated animal. Then the child was asked to make moves with the source so that the beam fell on other animals.

Of course, the target could be moved in any position or the light could be reflected from two or even three of the mirrors. The child was asked to specify the path the light took to hit the target. The co-ordination of adjustments on the part of the child has been of primary interest in this problem.

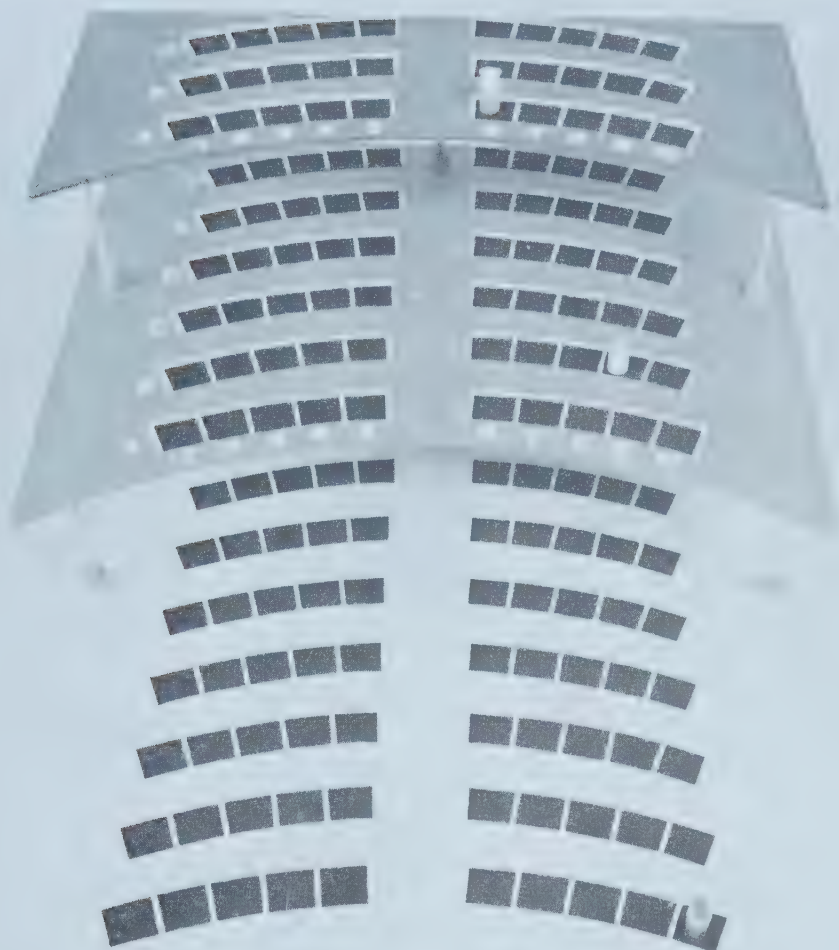


3. *Animal Groups.* The ability of the child to make equal groups either by measurement or partition could be determined by this problem. The layout is shown in the photograph.



It consists of a large board in which holes are drilled to make a grid. Grooved dowels are cut down so that they fit in the holes of the board. The fit is quite tight so the dowels stand up rigidly. Panels are cut to fit the grooves in the dowels. These panels can be inserted by the child to make cages for the animals. The photograph indicates how the cages are assembled. One basic problem was to give the child a set of the animals and get him to make enough cages so that a given number of animals would go in each cage (measurement division). Another was to give the child a set of animals and a certain number of cages. He was asked to distribute the animals among the cages so that each cage would have the same number of animals (partitive division). Of course, a great variety of problems of this type can be devised around this particular apparatus.

4. *Theatre Grid*. In order to help children think about moving systematically in three dimensions, the theatre grid was developed as shown in the photograph below. The floors are each of a different color but the rows and seats in the row are indicated on each floor by a letter and a numeral. The child was then given a toy man and a ticket. The object was to find the correct seat for the man as indicated on the ticket. The color of the symbols on the ticket indicated which floor, the letter of the row, and the numeral of the seat in the row.



5. *Circular Sequences.* Problems associated with sequences can be designed around the carousel shown in the photograph.

The carousel has 12 basic compartments but each compartment can be provided with a partition to make 24 compartments in all. The front of the carousel can be adjusted from a closed position to positions which reveal one, two, or three compartments at a time. The compartments can be charged from the back in any kind of sequence with any kind of objects. The simplest would be to alternate two kinds of objects, for example, toy elephants and toy airplanes.

The child can be given information about the sequence by opening the front and turning the carousel. We are particularly interested in how much information a child needs to predict with certainty what the next object would be. A whole array of interesting problems come out of this apparatus.

6. *Factor Platform.* The factors of compound numbers can be illustrated physically by equal piles of objects. The factor platform shown in the accompanying photograph was devised so that the mechanics of piling would be minimized. A certain number of blocks were fitted in a number of grooves in piles of differing heights. The child then predicts whether the blocks can be rearranged so there will be the same number of blocks in each





row and verifies his prediction. Various strategies for testing the prediction can be observed among young children.

Basic Research Procedures

The research referred to in the early part of this paper is concerned with creating good problems, developing protocols for use with the problems, recording the behavior of children while engaged in the problems, and finally, analysing the problem solving behaviors.

To accomplish this purpose, six problems and a set of six different but mathematically equivalent problems were developed. A sample of ninety children was selected, fifteen from each age level ages three to eight. Each child was scheduled for a problem solving session. There was no time limit on the session, but the child was presented with six different problem situations in the session. Most children completed the session in less than an hour and some took as little as twenty minutes.

While children were engaged in the problems, their behavior was recorded on $\frac{1}{2}$ -inch video tape (black and white). This is the only record that was kept. The behavior of all children in the sample produced 54 hours of taped data.

Using the Video Tape Recorder as a Data Gathering Device

There is one very obvious advantage to taped data. They represent a permanent record of the actual behaviors of the children. One doesn't run into the same problem of reliability in the interpretation of the data as is usual with other methods of observation. If there is a question, one can always return to the actual event simply by running the tape through again. However, it is time consuming to search through the mass of original data to check out some short sequence on tape. Regardless of the thought given to organizing and storing the taped information, retrieval is never easy and efficient.

At some point the information has to be taken from the tape and coded in a systematic way. The coding system can be worked out by observing the actual tapes and does not have to be done beforehand as would be the case if the data were coded while the behavior was going on. Those who are coding taped behaviors do not have to be so highly efficient in using the code from the beginning. Their skill and efficiency in using a code can be developed as they record. Reliability checks can be very quickly run and blinds can be easily applied. If a coder cannot learn to record the behaviors in a reliable way, he can be replaced with no serious harm to the research.

Another advantage to taped data is that one is readily able to detect flaws in consistency in the use of protocols. Slight variations in protocol may occur unnoticed if behaviors are being coded as they occur.

One has also to consider shortcomings of using taped data. The behavior of children when a camera is trained on them may not always be the same as it would be without the camera. Also there is a necessity to introduce extra people and apparatus into the situation which may distract the child. Fortunately in the research reported here, the problems were attractive enough to most children that such distractions were minimal. Another difficulty is that the behaviors have to occur in a relatively small space. Behavior cannot be faithfully filmed if the child is permitted to move about a great deal. It may be that for some problems, such movement would be desirable.

6 6 0 0 V₃ "Yes" R^②_T R^②_{↓2} L₃ R₃ L→R^①_↓ Lp(ea. bl.)_{c.2} R₃ L₃[→]
4 4 4 0

4 4 4 0 E B_T L₆^③ ₆ ₅ ₄ ₃ L₃^②_↓ L₂ R₆ L₅ L₄ L₃^①_↓ R₃
2 2 2 2 2 2 R₂ R₁ E₃

E R₅^⑥ R_↓^⑥ L_↓^⑤ V₃ "No" E "What happens?" V₃ "I gotta
2 2 2 2 2 0 (holds 2) have 3 more." Bm_{1,5}
3 2 2 2 3

Note: This example illustrates the format of the code. Detailed interpretation of the code will be included in the final report to Canada Council.

It has been necessary to devise a coding scheme for each problem individually. The code must be easily read and must include those behaviors which are important but must eliminate other behaviors that are not stimulated by the problem. Coding for some problems has been completed. Below is an incomplete sample of coded behavior of a seven year old child working with the factor platform (described earlier).

Interpretation of these codes has not proved too difficult and patterns of behavior can often be detected just by looking at the coded material. If ever a particular coding is found to provide an inadequate picture of what actually occurred, a revised system can be developed without loss of data. If on-the-spot coding is done and the code doesn't work properly, all the data recorded under the defective system would be lost.

Some Further Observations

Although analysis of the data from this project is far from complete, there are a few observations which can be made at this time.

1. With the use of the guidelines developed here, it is possible to create problems which stimulate a great deal of problem solving activity on the part of young children. Even though the mathematical process involved may not have been fully understood by these children, the activity they engaged in with the material was almost always directed to some kind of solution. Only one extremely shy child in the entire sample of ninety children could not be induced to react in the problem situations. The others tended to become so involved in problem solutions that they completely ignored the rather overwhelming laboratory set-up.

An encouraging aspect of the guidelines is that once they are familiar to a person with a reasonable mathematics background and an understanding of child development, there seems to be no real limit to the number of "good" problems that, given time, the person could create.

2. Even very young children display a surprising range of activities which have mathematical significance. Making groups of objects with the same number in each group, fitting corner to corner and edge to edge to make three dimensional shapes, displaying rather complex sequences, moving about confidently in a plane, copying complex geometrical solids are only a few of the activities readily engaged in by children in the whole age range.

3. Some seemingly simple procedures which are assumed to develop early fall far short of expectations. For example, only a few children at the upper age levels showed that they had mastered the partitioning process. (From a given set of objects, putting the same number of objects in each of a given number of subsets.) Also controlling the reflection of light on an object in a plane and predicting or explaining what happens when an object or a beam of light is reflected in a plane is apparently very difficult for young children.

Children can describe the properties of a set of discrete objects if they are permitted to look at all the elements in the set at once. However, if the information is given to them an element at a time in a particular order and repeatedly, many children are not able to induce and describe any common properties of the elements.

4. Many of the older children in the sample could have solved many of the problems by making a simple mathematical computation, but in these situations, hardly any children did so. They seemed to prefer solving the problem by making moves or transformations with or in the material provided in the problem. It seemed as if they viewed the physical solution as something completely independent of any mathematical computation. For example, children who knew that 19 has no divisors would keep trying to arrange a set of 19 blocks in a number of equal piles. Or children who knew that $15 \div 3 = 5$ would use a ferry boat to take three toy cars at a time (from a set of 15 toy cars) across a river and count the number of trips. The actual moves seemed to provide the kind of verification they needed.

J. MARTIN

The University of Alberta

Controlled vs. Natural Settings—Some Implications for Behavioral Analysis and Change in Classroom Situations

This paper considers five empirical and theoretical issues which have arisen from a comparison of behavioral analyses in controlled laboratory settings and similar analyses in naturalistic human environments. The empirical bases upon which some standard behavioral variables and processes rest are found to differ slightly from one situation to the other. Implications for the behavioral analysis of classroom environments are discussed. (Dr. Martin is a Lecturer in Educational Psychology, The University of Alberta.)

A prodigious amount of literature under the general rubrics of "reinforcement theory and its applications" or "behavioral modification" has been accumulated over the past two decades. More recently, the empirical foundation supplied by such work has been incorporated by Bugelski (1964), Becker, Engelmann and Thomas (1971) and others, and logically extended to the area of classroom learning and instruction. On the basis of the research to date, it is clear that the behavioral concepts and principles developed thus far can be applied to the teaching situation and can produce gratifying results. Statements of demonstrated relationships between environmental events and behavior such as "The longer the interval between an operant response and its reinforcing stimulus, the less the strengthening effect on responses of the same class" are of obvious importance to the classroom teacher. These statements are the *facts* of the science of psychology as generated by an experimental analysis of behavior. However, while the growth of a psychology of learning and instruction built upon the results of behavioral analyses in controlled, manipulated environments may legitimately be viewed as an important contribution to educational practice, the essential extrapolation of procedures and findings to natural classroom environments can be shown to involve certain empirical and theoretical difficulties.

During the past three years, the writer has been engaged in developing methodological procedures which permit a systematic analysis of

behavioral contingencies as they operate in non-manipulated small group situations—predominately student learning groups, and psychotherapy groups (*cf.* McLeish & Martin, 1975). Taken together, these methodological innovations make possible an exacting analysis of naturalistic behavioral data, in much the same manner to which behaviors are studied in controlled environments. A comparison of the results obtained in these two settings highlights some problems which must be resolved if behavioral science is to achieve the practical success in education which its evolution as a positive science would promise. Without being “pejorative” of present practices in the psychology of learning and instruction, the paper aims at alerting educators, psychologists, and educational researchers to difficulties which urgently demand further thought and analysis. Five such issues will be reviewed.

1. *Determining the functional properties of environmental stimuli.*

For the functional behaviorist meaning is synonymous with function. The meaning of any piece of activity is determined by the sets of environmental stimuli which precede and follow upon its emission. When a baseball player swings at a pitch, he does so because he has been reinforced in the past by an ensuing drive into deep center or some such place. If a subsequent environmental event increases the frequency with which a response is emitted, it acts as a reinforcing stimulus. Its reinforcing properties are disclosed by a study of the future emission pattern of the behavior upon which it follows. This is the only valid criterion for determining whether a stimulus is reinforcing in any particular situation.

Unfortunately, in the actual practice of behavior management in the classroom (*e.g.* Zimmerman & Zimmerman, 1962; Ward & Baker, 1968; O’Leary *et al.*, 1969) this criterion is often superceded by judgments based upon the “essential” positive or negative natures of particular stimuli. On a structural basis, certain stimuli such as “warm smiles,” “friendly nods,” praise, round wooden tokens, and merit points are assumed to be positive reinforcers for all individuals at all times.

Now while there may be some consensual or statistical validity for the view that certain stimuli are more positive in more situations than are other stimuli, this is an extremely dangerous assumption. The mere fact that most of those who are subjected to behavior management programs have not managed to acquire conventional social responses to standard stimuli in their environments, should warn us that functional relationships other than conventional ones may be operating in many of these cases.

For example, in an experiment with psychotherapy outpatients as subjects (McLeish & Martin, 1975), it was discovered that aggressive interruptive behaviors were positively reinforcing for the general vocal behavior of two female subjects, whereas emphatic therapeutic responses produced behavioral frequencies indicative of the operation of punishment. Because of this and much other similar evidence, it would seem that there is no stimulus which serves a universal function, and that behavior management programs which employ such concepts, either implicitly or explicitly are doomed to failure in the specific case. No classroom program will be optimally successful unless the stimuli which are manipulated are appropriate for individual students.

2. Changes in the functional properties of stimuli over time.

The second issue is in some ways a corollary of the first. In some classroom management programs the implicit assumption is made that the functional properties of discriminative and reinforcing stimuli remain static over the temporal course of the modification program. If a particular stimulus acts as a positive reinforcer for a particular response at one point (A) in time, this does not necessarily imply that the same operant relationship will remain intact at a future point (B). Further, the time interval between points A and B need not be long. Our empirical results show that as little as five minutes (and sometimes less) may be sufficient to shape-up such a functional change.

Perhaps the best example of this phenomenon is taken from an analysis of the contingencies of reinforcement which operated during a student learning group composed of freshman students at the University of Alberta (McLeish & Martin, 1975). In the first two minutes of the seminar, the instructor's "emphatic" cephalic nods attached themselves as positive reinforcing stimuli to particular verbal responses of student A. However, this ideal theoretical pattern was eclipsed in the succeeding four minutes of interaction so that as the instructor's nods increased dramatically in their frequency of emission, they became functionally aversive, either attaching to student A's responses as punishing stimuli, or to the responses of other students as negative reinforcers in contingencies of avoidance.

Any naturalistic situation is characterized by a dynamic chaining of behavioral patterns. In the classroom, a response in one contingency for one student may act as a reinforcer for a different response by a second student and may also be caught up in a discriminative stimulus complex for a series of future responses. The competitive inhibitions and associations which result from this dynamic vortex seem to shape and modify specific stimulus properties even within relatively short time intervals.

3. Stimulus deprivation vs. extinction

Perhaps the most obvious difficulty that is encountered when behavioral analyses and techniques are extrapolated from a laboratory situation to a naturalistic human environment involves the inadequacy of traditionally established behavioral processes such as reinforcement, punishment and extinction in explaining all fluctuations in the emission rates of "social" behaviors.

McLeish and Martin (1975) found it necessary to consider a process which they have called "stimulus deprivation." Although this is certainly not an original term, its use in the present context is unique and specific.

If it were possible to hold a human social environment constant in the same manner as laboratory environments are regulated, the process of extinction would be sufficient to explain all systematic decreases in emission rates. However, because any natural social situation is dynamic in its stimulating qualities, operant responses may decrease in frequency *not* because they cease to be adequately reinforced, but because they are not subjected to the discriminative stimulus situations which come to exert control over their emission. Each time an individual in a classroom speaks or moves, his action alters the social environment in some way for other par-

ticipating individuals, and indeed for himself as well. Potential discriminative stimulus situations are thus constantly changing.

Even though a particular student is consistently reinforced for his speaking behaviors, the classroom situation is not always propitious for his contributions, and periodic decrements in the rate of emission of his speaking behavior will be witnessed. No withholding of reinforcement has occurred. What has occurred is that preceding discriminative stimuli have been withdrawn. The withdrawal of a discriminative stimulus is thus a common process in the human environment, and as such is labelled stimulus deprivation.

As a new era of sophistication in behavior technology and programming unfolds, it may be necessary to consider additional processes which have not been historically established by the experimental works of Skinner and company.

4. "Withdrawing" or "withholding"?

A fourth difficulty is that even some of the standard behavioral processes may require further clarification if they are to have optimal explanatory power in classroom situations. The concepts of punishment and extinction can be particularly troublesome.

Skinner (1953) defines two kinds of punishment—the *withdrawal* of a positive stimulus (type 1) and the presentation of an aversive stimulus (type 2). Extinction involves the *withholding* of a reinforcing stimulus. The distinction between extinction and punishment of the first type is reflected in the difference between *withholding* and *withdrawing*. When a rat in a Skinner box presses the bar and the magazine fails to click and a food pellet (the positive reinforcing stimulus) fails to drop, we say that positive reinforcement has been *withheld*. Our use of the term withheld is semantically justified by the fact that the rat, prior to this occurrence, had received a food pellet each time he depressed the bar. When a child is playing with a toy and we take the toy away from him, we *withdraw* the positive reinforcing stimulus. The difference in terminology is justified by the fact that in the case of the rat, the *withheld* food pellet is not directly perceived by the behaving organism, whereas, in the case of the child, the toy is actually in his possession and directly perceived by the senses before being removed. The withholding of the pellet in the first case leads to extinction. The number of "withholdings" is sufficient to overcome the response pattern established by whatever schedule of reinforcement was previously in effect. The withdrawing of the toy in the second case counts as punishment—type 1.

The two examples thus considered present no theoretical difficulty. However it is possible to provide numerous examples from classroom situations which defy this logical exclusivity. When a student's behavior is being positively reinforced by a smile from the instructor, and the instructor ceases to smile, is this punishment or extinction? If the smile is "removed" precisely between two successive emissions of the student's response, it would seem to count as extinction. This is because the smile is withheld after a response which has been previously reinforced. However, in dealing with a time interval which may be as small as one second, the distinction based on such arguments seems rather forced.

The student's contact with the positive reinforcing stimulus (the instructor's smile) is somewhat different from the rat's contact with the food pellet. The individual can actually observe the change in the instructor's expression as his smile disappears. The rat however, does not observe the mechanical alterations which affect the functioning of the food magazine. If the instructor emits a continuing smile during several emissions of the student's response and then alters his countenance for a subsequent emission, is this withdrawal or withholding? Clearly, more attention to this problem is warranted.

Extinction differs from punishment of the second type in a more obvious manner. In extinction a negative reinforcing stimulus may be withheld in order to curtail a contingency of negative reinforcement, but at no time does the presentation of an aversive stimulus count as extinction.

	Punishment Type I	Punishment Type II	Extinction	Stimulus Deprivation
Definition	withdrawal of +Re	presentation of aversive stimulus	withholding of +Re or -Re	withdrawal of discriminative stimulus

Another source of uncertainty is the chronological order of stimulus, response and reinforcer. If punishment is the withdrawal of a positive stimulus, stimulus deprivation would appear to be a form of punishment. Skinner, however, illustrated punishment only with examples of the withdrawal of, or presentation of, stimuli which succeed the response.

5. *Operant vs. classical conditioning*

Still another theoretical issue is concerned with the importance of preceding stimuli in controlling behavior. Skinner has established a science of behavior which is primarily concerned with the control which subsequent events exercise over responses. The effect of this is that rigid demarcation lines have been drawn between (so-called) "operant" conditioning and "classical," Pavlovian or "respondent" conditioning. This distinction has been unfortunate. It has served to limit unduly the orientation of research workers. Experiments in operant conditioning are considered by many to be valid only if they adhere to the demarcation lines. Skinner himself, while recognizing that respondent conditioning does operate in certain instances, and that even in "operant" conditioning discriminative stimuli exercise some control over behavior, tends to bypass this problem. He is more interested in carrying out an in-depth analysis of reinforcers and their effects on behavior. Nevertheless, the decision to do so seems to be arbitrary. If two events occur in a chain of other events it may be rather limiting to consider these events in terms that one event precedes the other, while ignoring the total chain of causation in which they are embedded. Skinner looks only at *half* of the conditioning process: the other half is supplied by Pavlov.

In truth, Skinner's concern with reinforcing stimuli to the exclusion of eliciting stimuli seems more methodologically pragmatic than theoretical. While it is difficult to quantify all the physical properties of an experimen-

tal space, especially in terms of temporal distances between stimuli and responses, it is less difficult to quantify the physical properties of an experimentally manipulated reinforcing event and the time intervals between responses and subsequent stimuli. This produces nice “clean” experimentation in the animal laboratory. But the clarity and cleanness are illusory. In naturalistic human settings, in the absence of experimental manipulations of variables, the research worker finds that reinforcing stimuli are *just* as difficult to recognize as are preceding (discriminative) stimuli. The pragmatic advantage of concentrating solely on subsequent events in terms of controlling behavior is lost.

There seems, therefore, to be no valid reason for limiting the scope of naturalistic investigation of classroom behaviors to *operant* phenomena, as defined by Skinner. Indeed, an overall picture of functional connections between behaviors, and between behaviors and the environment, would be enriched by a closer scrutiny of the physical aspects of the stimulus situation to which subjects are responding. The rigorous scientific work of Skinner, supplemented by the constructs of Ivan Pavlov and his model of the intact behaving organism, offers a sound base on which future work might rest.

Concluding Statement

The training of teachers in the use of behavioral procedures offers great promise for the efficient management of student learning behavior. However, behavioral scientists, in the face of severe (and all too often “emotionally-laden”) criticism from those oriented towards the concepts of phenomenology, humanism, existentialism and psycho-analysis are often prone to overstate their case, and to go beyond supportive evidences. As Skinner (1972) states:

. . . we are on the verge of a new pedagogy—in which the teacher will emerge as a skilled behavioral engineer. He will be able to analyze the contingencies which arise in his classes, and design and set up improved versions. He will know what is to be done and will have the satisfaction of knowing that he has done it (p. 233).

Such a pedagogy is indeed a possibility of the future, but many difficulties in the present state of behavioral knowledge must be resolved before the veracity of Skinner’s “apocalyptic vision” can be reasonably assessed.

Behavioral Analysis and Change in Classroom Situations

References

- Becker, W. D., Engleman, S., & Thomas, D. R. *Teaching: A course in applied psychology*. Chicago: Science Research Associates, Inc., 1971.
- Bugelski, B. R. *The psychology of learning applied to teaching*. New York: The Bobbs-Merrill Company, Inc. 1964.
- McLeish, J., & Martin, J. F. Verbal behavior: An experimental analysis. *Journal of General Psychology*. 1975, in press.
- O'Leary, K. D., Becker, W. D., Evans, M. B., & Saudargas, R. A. A token reinforcement program in a public school: A replication and systematic analysis. *Journal of Applied Behavior Analysis*, 1969, 2, 3-13.
- Skinner, B. F. *Science and human behavior*. New York: The MacMillan Company, 1953.
- Skinner, B. F. *Cumulative record*. (3rd ed.) New York: Appleton-Century-Crofts, Inc., 1972.
- Ward, M. H., & Baker, B. L. Reinforcement therapy in the classroom. *Journal of Applied Behavior Analysis*, 1968, 1, 323-328.
- Zimmerman, E. H., & Zimmerman, J. The alteration of behavior in a special classroom situation. *Journal of the Experimental Analysis of Behavior*, 1962, 5, 59-60.

S. VERMA

and

D. L. PETERS

The Pennsylvania State University

Day Care Teacher Practices and Beliefs

A self-administered rating scale and an observational schedule were developed to assess the beliefs and behavior of early childhood teachers in relation to constructs derived from Piagetian cognitive developmental theory and operant learning theory. Application was made to existent Piagetian and operant model early education programs. The results indicated both measures satisfactorily differentiated among the programs in a manner consistent with theory. The instruments were then applied in 38 day care centers. The results indicate that although consistent and meaningful differences are found among teachers, a wide range of internal inconsistencies exist between a teacher's stated beliefs and her actual practices. (Dr. Verma is with the Department of Education, Capitol Campus, The Pennsylvania State University at Middletown; Dr. Peters is Associate Professor of Human Development, Division of Individual and Family Studies, The Pennsylvania State University at University Park.)

During the last ten years great efforts have been made to increase the quantity and quality of early education programs. As part of this effort day care has become a topic of major concern and the number of day care facilities has increased markedly. Yet many questions concerning the process or transactions occurring within day care centers remain (Grotberg, Chapman, & Lazar, 1971). Our lack of information is the more remarkable because the analysis of process variables generally is considered an integral part of evaluation (Stake, 1970; Provus, 1969; Scriven, 1967) and quantitative descriptive studies are considered an important first step in understanding the teaching process (Rosenshine & Furst, 1973).

In order to adequately evaluate day care programs it is important to develop measurement procedures which differentiate existing programs from one another in terms of both the observable transactions that occur and the underlying assumptions which constitute the rationale for the

program. Measures of this type would permit the isolation of variables which have potential differential effects on the children enrolled. The purpose of this study was the development of appropriate and theoretically relevant measures for the naturalistic observation of teacher/child interaction patterns within day care settings and for assessing the beliefs or attitudes held by the observed teachers.

A recent extensive review of the literature uncovered only twenty-six data based observational studies of preschool teacher behavior published during the period 1963-1973 (Gordon & Jester, 1973). Of those, only one study by Prescott, Jones and Kritchevsky (1967) directly addressed interaction patterns within day care settings. Several recent studies have added to the day care list (Jambor, 1973; Green, Hollick, Knowles, Vanderkar & Winter, 1973). Yet, none of the measures reviewed fully meets the criteria we believe should be imposed when transaction measures are used within an evaluation framework (Stake, 1967). In addition to the usual psychometric requirements, an appropriate system should:

1. have a basis in some rationale or theory appropriate to early education (*i.e.* there should be some *a priori* reason for thinking the observed behaviors are important to early childhood programs);
2. provide a means for determining the internal consistency of behaviors with each other and with the program rationale (Stake, 1967);
3. be capable of distinguishing among programs in expected ways; and
4. provide an empirical basis for demonstrating that the variables selected for observation are related to child outcomes.

A brief review of prior observational systems will help to make the point. Prescott *et al.* (1967) developed a category system based on the earlier observations of Reichenberg-Hackett (1962). The resulting observation procedure included such behavioral categories as teacher direction, guidance, restriction, development of verbal skills, neutral activities and non-communications. Each category was elaborated with a system of sub-categories. While useful for describing some aspects of the day care environment as it existed in the fifty centers observed, the procedure offers only limited data for assessing the appropriateness and effectiveness of the teacher behavior observed.

Jambor (1973) extended and modified the Prescott and Jones categories to provide a more specific rationale for judging the appropriateness and consistency of teacher behavior within day care. He reorganized the behavioral categories into three broad role model headings suggested by Katz (1970): the maternal role, the therapeutic role and the instructional role. Both Katz and Jambor argue that behavior consistent with each of these roles is required to fully meet the needs of the day care child. Jambor then went on to observe the behavior of 10 nursery school and 10 day care teachers to determine the dispersion of their behavior across the three roles. While he uncovered large individual differences, he found no significant group differences between the day care and nursery school teacher groups. Both engaged in many more instructional and maternal interactions than therapeutic ones. Little relationship was found between any of the teaching roles and a variety of classroom and teacher characteristics.

Though based on practical logic and providing a means for judging the

internal consistency of behaviors, Jambor's categories lack both a theoretical foundation and prescriptive power as they have been developed to date. The data produced are not immediately suggestive either for day care program improvement or for revision of training programs for day care personnel. No attempt was made to relate any of the behaviors to child outcomes.

The *Teacher Practices Observation Record* (TPOR) developed by Brown (1968) was also designed to assess the relation of teacher behaviors to a standard, in this case Dewey's experimentalist position. The TPOR consists of 62 items, half of which represent the experimentalist position and half do not. Soar (Soar & Soar, 1972), using a modification of the TPOR for comparison among Follow-through programs, factor analyzed his data and identified several factors which identify teacher behavior with reference to program type (e.g. Bereiter-Englemann *vs.* British Infant School).

Although one can question the relevance of Dewey's experimentalism to the day care setting, Soar's results do suggest that programs may be readily differentiated, and that training a teacher in a particular program model tends to lead to more uniform behavior. The procedure does not, however, indicate if the uniformity is consistent with the specific rationale of the program. Soar and Soar (1972) do find that several behavior factors relate to pupil achievement.

An alternative direction is suggested by the research of Harvey, White, Prather and Hoffmeister (1966). These researchers were interested in assessing the relationship between a teacher's own beliefs and her classroom behavior. Using a sentence completion procedure the beliefs of 168 preschool teachers were assessed. Thirty teachers representing three different belief groups (high abstract, low abstract and high concrete) were then observed using a twenty-six dimension rating scale. The rating scale included such dimensions as anxiety, punitiveness, warmth, flexibility and so forth. The comparison of the three belief groups' ratings indicated there was a significant relationship between beliefs and the ratings of behavior. However, there is no attempt to relate the teacher's beliefs to any generic early childhood rationale, nor is there an attempt to relate either beliefs or behavior to child outcomes.

The studies reported here constitute a preliminary attempt to further identify variables which have potential for discriminating among early childhood programs and to develop transactional measures which meet the criteria we have imposed.

Development of Measures

Theoretical Framework

The preceding literature review suggested that an appropriate point of attack for assessing the sources of differences across programs would be teacher behaviors *and* the belief system upon which they are based. That is, since program rationales differ, if one is to assess congruence of teacher behavior with the program rationale rather than some external standard, then the rationale (beliefs) need to be measured as well as the behavior. Further, if the variables isolated for study are to have potential impact on the children enrolled, it was felt that they should be derived from the

applications of developmental theory: theory relevant for early childhood program development.

The theories of Piaget and Skinner are vastly different but both have contributed in a major way to the development of early childhood programs during the past ten years. Both have been cited extensively in the early childhood literature and have been given a place in the rationale for most, if not all, recent early childhood program developments. For this reason they were selected as the basis for instrument development.

Piagetian and Operant theory formulate different assumptions about children, development and learning (DeVries, 1974). A careful review of the literature related to each suggested ten dimensions of potential value in that the applications derived from the two theories were quite different (Verma, 1973) (See Table 1), and belief statements and observable behavior categories could be generated from each. Five of the dimensions were selected for study based on an estimate of their importance to the theory and the practicality of their use. A number of belief statements and behavior categories were generated for each dimension using items from Stern and Gordon (1967) and from Brown (1968) whenever possible.

TABLE 1
THEORETICAL ASSUMPTIONS OF TWO THEORIES

	Piaget's Theory	Operant Theory
Children	*1. Active	1. Passive
	*2. Qualitatively unlike adults	2. Qualitatively like adults (adult miniatures)
Development	3. Qualitative	3. Quantitative
	4. Interaction between child and environment	4. External environment
Learning	*5. Intrinsic motivation	5. Extrinsic motivation
	6. Based on sensory education	6. Based on language overt and covert verbal labeling
	7. Stage dependent	7. Knowledge and familiarity with task or tasks similar to it
	*8. Based on massive general type of experience	8. Based on specific training
	*9. Process approach	9. Product approach
	10. Irreversible	10. Reversible
	(invariant)	

* Dimensions retained in final instruments.

Teacher Belief Rating Scale

A preliminary version of the Teacher Belief Rating Scale (TBRS) was developed by converting the 6 items generated from each of the five selected dimensions to a Likert-type scale. Piagetian and Operant items were randomly ordered and rating order was reversed intermittently to avoid the formulation of response sets. Two scores per respondent were attained by summing Operant and Piagetian items separately (with appropriate reversals).

The scales were then submitted to the program consultant or subject matter specialist, the head teacher and the assisting teacher of the Piagetian and Operant based laboratory early education programs of the University. As such these two groups ($N=3$ per group) were known to differ on the basic constructs being measured. Analysis of variance revealed the effects of scale to be significant ($F = 20.5$; df 1, 8; $p < .05$). The group by scale interaction was also significant ($F = 36.8$; df 1,8; $p < .05$). This interaction was considered particularly crucial for determining whether the TBRS was performing properly.

Suggestions of the experts and of eight day care teachers concerning readability and applicability resulted in the deletion of six items. The final version of the TBRS therefore consisted of 24 items (12 representing operant beliefs and 12 representing Piagetian beliefs). This version was subsequently given to eleven early childhood personnel known to be affiliated with Operant, Piagetian or other programs. Analysis again indicated as scale main effect ($F=19.8$; df 1,8; $p < .05$) and a scale by group interaction ($F = 34.5$; df 3,8; $p < .05$).

Teacher Practices Observation Form

The Teacher Practices Observation Form (TPOF) was developed by formulating observable behavioral categories each of which correspond to the items on the belief scale. As such, the TPOF underwent a parallel development to the belief scale. The initial observation form consisted of 30 items (15 Piagetian and 15 Operant—six from each of the five dimensions). After successive trials in three early childhood education programs, six categories were deleted and the remaining 24 were collapsed into 12 numbered items with a Piagetian, an operant, and a not observed alternative under each. Thus the observer was required to go down the list, checking one of the alternatives for each item immediately following a 15 second observation segment.

The TPOF then was used to collect data from the Piagetian and Operant laboratory early education programs using the procedures indicated above. Since the programs observed were training programs, observations were collected of the behavior of both the teachers and the student teachers. Average scores were calculated by dividing the total checks per item in each program with the total time observed. This procedure provided a rough rate measure for each of the teacher behaviors. Directionality was hypothesized on the basis of the scale from which a particular item was drawn. A sign test was used to determine if the relationship between rates of observed practices were consistent with the directions hypothesized by theory. Twenty of the 24 items showed differences in the predicted directions. Two items were not observed in either program and two items showed a reversal of the predicted direction. In general, the results indicated

that the observation form was successful in differentiating the Piagetian teachers' behavior from Operant teachers' behavior even when some of the teachers are not yet fully trained.

In another study (Cohen, Peters & Willis, 1974), the behaviors of student teachers were observed in each of four early education programs using the TPOF. A total of 120 observations in each room was made over a two month period. The observed behaviors for each scale were summed across student teachers by program. The total observed frequency of Piagetian and Operant behaviors are presented in Table 2. These data tend to further confirm the validity of the TPOF and its ability to differentiate across early education programs.

TABLE 2
OBSERVED FREQUENCY OF PIAGET AND OPERANT BEHAVIOR
FOUR PROGRAMS

Behavior	Program Type			
	Operant	Piagetian	Responsive Environment	Day Care
Piaget Behavior	64	472	433	361
Operant Behavior	619	96	161	206

Reliability of the Measures

Reliability of TBRS. Kuder-Richardson Formula 20 internal consistency reliabilities were estimated for the TBRS utilizing data from the total study sample ($N=38$). Two methods were used. In method one each agreement statement (slightly agree, moderately agree, strongly agree) were coded as "2" for each scale separately. Disagreement statements were coded as "1". The reliabilities thus calculated were .44 and .58 for the Piaget and Operant scales respectively. The intercorrelation of the two scales was .13 (non-significant).

In the second method Kuder-Richardson Formula 21 internal consistency reliabilities were calculated using the full 6 point scale range. In this case the reliabilities were .56 and .66 for the Piaget and Operant scales respectively. The intercorrelation of the two scales was -.04 (non-significant).

Inter-observer agreement for TPOF. 68 observations were used to pretest the final Teacher Practice Observation Form. Two observers simultaneously observed the same teacher 68 times (investigator made 68 observations and observer A and B made 42 and 26 observations respectively). The percentage of agreement between the observers was calculated.

The mean agreement across observers ranged from .72 to .96 and across items from .60 to 1.00. The overall mean agreement was .83.

Application to Day Care

Sample

The sample for the study consisted of 38 presently employed teachers in day care centers in Pennsylvania. The centers were selected from existing

lists using the criteria that the programs should be geographically representative and have more than 15 children currently enrolled.

Procedures

Two trained observers were used in the study, each visiting approximately half the programs. During the visit the head teacher was observed between 9:30 a.m. and 11:30 a.m. The observations were made during three half-hour periods with a ten minute break between them. Twenty observations were made during each half hour period using a point-sampling procedure. This yielded a total of sixty observations for each teacher. Throughout the visit the observers remained as unobtrusive as possible, though the teachers naturally knew they were being observed. After the observations were completed the teachers were requested to respond to the Teacher Belief Rating Scale.

Results

The results of the field study indicate that the scores on the belief scale and the frequencies of observed behavior were normally distributed across the sample. No clear Piagetian groups or operant groups of teachers could be discerned on the basis of either beliefs or practices. Spearman rank correlations were calculated for the total scores on Piagetian beliefs and practices, and for total operant beliefs and practices. These data indicate that only a weak relationship exists. Teachers who rank high in Piagetian beliefs also tend to rank high in Piagetian practices ($r=.30$, $p < .05$). Although a positive relationship is indicated for the correspondence of ranks on operant beliefs and practices the results do not reach acceptable levels of significance ($r=.10$, $p > .05$). A significant negative correlation was found between the ranking of Piagetian practices and Operant practices as might be expected from the nature of the construction of the observation instrument ($r=-.36$, $p < .05$).

In general the results indicate that the day care teachers agree significantly more with Piagetian beliefs ($\bar{X}=51.9$, $sd=8.57$) than with operant beliefs ($\bar{X}=42.1$, $sd=10.0$) ($t=4.69$, $p < .05$), but behave in ways more consistent with operant theory ($\bar{X}=113.3$, $sd=28.9$) than with Piagetian theory ($\bar{X}=71.7$, $sd=35.7$) ($t=5.77$, $p < .05$).

Both these findings suggest that there is little correspondence between teachers' professed beliefs and observed practices.

To adequately assess the internal consistency of each teacher's beliefs and practices simple correlation coefficients of Piagetian beliefs and practices and Operant beliefs and practices were computed for each teacher. The results indicated that only two of the thirty-eight teachers had internally consistent beliefs and practices: one Piagetian and one Operant.

Discussion

It is not surprising that no clear Piagetian or Operant groups of teachers were found in operating day care programs. Nor is it surprising that only a weak relationship between behavior and beliefs was uncovered. Such findings have been reported before (Soar & Soar, 1972; Brown, 1968). Further, little effort has been made to implement theoretically pure or internally consistent programs in day care. Rather each teacher has attempted to put together her program in the way that appears to work best for her. When the range of alternatives available in materials and methods is con-

sidered, it is not surprising that the decisions and choices are difficult to make.

Perhaps the teachers of this sample chose the operant practices more often because they involve teaching techniques which are easier to learn and take less time to implement. Other studies have suggested this might be the case (Bissell, 1971; Miller *et al.*, 1970). Yet the choices made have led to an unnatural split between theory and practice. Only two teachers had beliefs and practices that were internally consistent. It appears that theory is little applied as either a policy for guiding practice or as an after-the-fact explanation for practice. This results in both the loss to practice of what is so promising in theory and the loss to theory of what is learned in practice.

The importance of the relationship between theory and practice has been stressed recently (DeVries, 1974; Parker, 1974; Peters, Cohen & McNichol, 1974; Willis, Cohen & Clement, 1974; EPIE, 1972). An explicit and solidly based theoretical stance is seen as an important ingredient for the improvement of early childhood programming. Theory provides a guideline for planning and acting in consistent ways. Through recourse to theory the teacher can derive program goals, set priorities, develop appropriate teaching methods, select materials, deductively arrive at solutions to everyday problems, and predict appropriate responses to future situations. Operating without theory is coping; i.e., dealing with each situation as it arises.

Both the lack of a strong relationship between theory and practice and the lack of internal consistency in either, suggest the need for an increased emphasis on theory in the training of day care personnel. Current emphasis on skill training should not be divorced from the rationale underlying the selection and use of specific skills in given circumstances. Day care personnel are in need of the "whys" as well as the "how to's" if they are to function independently in planning and implementing quality day care programs. A broad range of competencies must be considered if teachers are to be thoroughly trained (Peters & Dorman, 1974).

Conclusions

The studies reported here were designed to further identify variables which have potential for discriminating among early childhood programs and to develop transactional measures which relate theory and practice. The results of the studies conducted to date indicate that there are discernibly different implications for practice that can be derived from Piagetian theory and Operant theory. Both theories have relevance for early education and potential for the development of measures which can differentiate among program practices. From the theoretical framework provided by the theories two measures were developed—The Teacher Belief Rating Scale and the Teacher Practices Observation Form. When applied to programs known to differ on their theoretical bases, both measures produced reliable data which differentiated among the programs in meaningful ways. When applied to the day to day operations of day care, theoretical purity was found neither in beliefs nor in practices.

Our studies are not yet complete. We have not yet met our own fourth criterion. That is, the relationship between the teacher behavior variables

selected for study and appropriate child outcome variables has not yet been determined. However, the progress made to date suggests that transaction or process measures for cross-program and within program evaluation can be derived by the steps we have followed. The TBRs and TBOFs appear to have reliability, validity and utility in day care program analysis. The results obtained from the major field test are suggestive for both future research and for the training of day care personnel. However, the final decision as to their importance awaits empirical assessment of the relationship between these teacher variables and changes in children.

This report is adapted from a dissertation conducted as partial fulfillment of the requirements for the Ph.D., College of Human Development, The Pennsylvania State University. The research was partially supported by the Department of Public Welfare of the Commonwealth of Pennsylvania. The opinions expressed are those of the authors and do not necessarily express those of the sponsoring agency.

References

- Bissell, J. S. *Implementation of planned variation in head start*. Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Child Development, April, 1971, ED 052 845.
- Brown, B. B. *The experimental mind in education*. New York: Harper, 1968.
- Cohen, A. S., Peters, D. L., & Willis, S. L. The effects of early childhood education on student teaching on program preferences, beliefs, and behaviors. (In preparation, 1974.)
- DeVries, R. Theory in educational practice. In R. Colvin & E. Zaffiro (Eds.), *Preschool education: A handbook for the training of early childhood educators*. New York: Springer, 1974, 3-40.
- Educational Products Information Exchange. *Early Childhood Education: How to select and evaluate materials*. Educational Product Report #42 New York: EPIE, 1972.
- Gordon, L., & Jester, R. Techniques of observing teaching in early childhood and outcomes of particular procedures. In R. Travers (Ed.) *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Green, R., Hollick, R., Knowles, K., Vanderkar, C., & Winter, M. The Pennsylvania day care study survey questionnaire and observation reports. Technical Report #9. Center for Human Services Development, The Pennsylvania State University, 1972.
- Grotberg, E., Chapman, J., & Lazar, J. A review of the present status and future needs in day care research. (A working paper prepared for the Interagency Panel on Early Childhood Education Research and Development: Office of Child Development, H.E.W., Washington, D.C., November, 1971.
- Harvey, O., White, B., Prather, M., & Hoffmeister, J. Teachers' belief systems and preschool atmospheres. *Journal of Educational Psychology*, 1966, 57, 373-381.
- Jambor, T. Instructional, maternal and therapeutic role behavior of day care and nursery school teachers. Unpublished doctoral dissertation, The Pennsylvania State University, 1973.
- Katz, L. Teaching in pre-schools: Roles and goals. *Children*, 1970, 17, 42-48.
- Miller, L. B. et al. *Experimental examination of head start curricula: A comparison of current approaches*. Progress report number 5. Research grant #OG-1199. Office of Economic Opportunity, Louisville, Ky.: University of Louisville, January, 1970 (a), ED 041 617.
- Parker, R. K. Theory in early education curricula. In R. Colvin & E. Zaffiro (Eds.), *Preschool education: A handbook for the training of early childhood educators*. New York: Springer, 1974, 41-74.

Day Care Teacher Practices and Beliefs

- Peters, D. L., & Dorman, L. Program planning and implementation. In R. Colvin & E. Zaffiro (Eds.) *Preschool education: A handbook for the training of early childhood educators*. New York: Springer, 1974, 115-137.
- Peters, D. L., Cohen, A. S., & McNichol, M. The training and certification of early childhood personnel. *Child Care Quarterly*, 1974, 3, 39-53.
- Prescott, E., Jones, E., & Kritchevsky, S. Group day care as a child-rearing environment. Pacific Oaks College, Pasadena, California, 1967.
- Provus, M. Evaluation of ongoing programs in the public school systems. In R. Tyler (Ed.) *Educational evaluation: New roles, new means*. (Sixty-eighth yearbook of the National Society for the Study of Education, Part II.) Chicago: University of Chicago Press, 1969.
- Reichenberg-Hackett, W. Practices, attitudes and values in nursery group education. *Psychological Reports*, 1962, 10, 151-172.
- Rosenshine, B., & Furst, N. The use of direct observation to study teaching. In R. Travers (Ed.) *Second handbook of research on teaching*. Chicago: Rand McNally, 1973.
- Scriven, M. *The methodology of evaluation: Perspectives on curriculum evaluation*. (AERA Monograph Series on Curriculum Evaluation, No. 1.) Chicago: Rand McNally, 1967.
- Soar, R., & Soar, R. An empirical analysis of selected Follow Through programs: An example of a process approach to evaluation. In I. Gordon (Ed.) *Early childhood education*. (The Seventy-first Yearbook of the National Society for the Study of Education). Chicago: University of Chicago Press, 1972.
- Stake, R. The countenance of educational evaluation. *Teachers College Record*, 1967, 68, 523-540.
- Stake, R. Objectives, priorities and other judgment data. *Review of Educational Research*, 1970, 40, 181-212.
- Stern, V., & Gordon, A. *Development of observation procedures for assessing preschool classroom environment*. Research Division Bank Street College of Education. New York, 1967.
- Verma, S. A study of the relationship between day care teachers' beliefs and teaching practices. Unpublished doctoral dissertation. The Pennsylvania State University, 1973.
- Willis, S. L., Cohen, A. S., & Clement, J. W. Formative evaluation in a cognitive developmental program for young children. Paper presented at the Fourth Annual Invitational, Interdisciplinary Conference: Piagetian theory and its implications for the helping professions, February, 1974.

ANNE D. FORESTER

University of Victoria

Learning the Language of Reading —An Exploratory Study

The general aim of the study was to gain insights into the thought processes and learning strategies of the beginning reader. It is based on the premise that since reading is a language task, the process of learning to read may show parallels to the process of learning to speak. Accordingly, the study was modelled after an observational study on the acquisition of oral language which yielded significant insights into the manner in which infants process oral language. It served as a pilot study for more extensive observations which have already confirmed some of the earlier findings.

Analysis of data collected during observations in a first-grade classroom suggest that there is a similarity between the processes of learning oral language and learning to read. The beginning reader, like the child learning to speak, appears to generate rules of his own to deal with the language input. Oral practice and feedback from the teacher guide learning more than the assimilation and application of rules presented in the classroom. Observations suggest that knowing the rules about sounds, letters, or endings follows rather than precedes the ability to manipulate these elements in the printed text. The child learns to sound out words once he has learned to read. (Mrs. Forester is a graduate student in Educational Psychology, the University of Victoria, British Columbia.)

The processes of learning to read and learning a language have a number of common elements: sounds and symbols have to be matched, new patterns of communication assimilated and tested for accuracy and semantic value. In both cases, the learner is faced with the need to assimilate and apply a multitude of rules in order to gain a degree of fluency. A child learns his native language without the benefit of formal instruction, seemingly without effort, and certainly without being apprised of the rules which govern that language. The child himself abstracts the patterns and regularities, and learns to transform them into new and sensible utterances. The beginning reader is faced with much the same task, that of learning to manipulate the rules of written communication, "making

sense" out of the language of reading. But in his case, rules and explanations are provided to help him learn.

In recent years, a number of studies have shed light on the processes involved in the acquisition of language and in view of the similarities referred to above, similar studies on the acquisition of reading appeared promising. A clearer understanding of the processes involved in learning to read might reveal to what extent the beginning reader—like the child acquiring language—is generating his own rules. If parallels would be found to exist between the two processes, such information could have a great deal of bearing on the type and number of rules provided in reading instruction and the timing at which they should be introduced.

The study here described¹ was undertaken in the hope that naturalistic observations in the classroom might provide the same type of data for analysis as that collected in the homes of infants learning to speak, in order to gain an insight into the thought- and learning-processes of the beginning reader. It was proposed as a small-scale replication of the study undertaken by Roger Brown and Ursula Bellugi (1964) on the child's acquisition of syntax, and has served as a pilot model for more extensive observations now under way with the use of VTR equipment which are corroborating and expanding some of the findings reported here.

In their study on the acquisition of syntax, Brown and Bellugi used tape recordings and direct observations of the verbal interactions between the subjects they studied and the subjects' mothers. The material thus obtained was then transcribed and analyzed to determine what patterns or processes could be inferred. Three processes were isolated: Imitation and Reduction, Imitation with Expansion, and Induction of the Latent Structures. Analysis of the latter process in particular stressed that inner processes of structuring more than imitation of adult patterns accounted for the acquisition of the more complex syntactic structures. It is described as "the very intricate simultaneous differentiation and integration that... is more reminiscent of the biological development of an embryo than it is of the acquisition of a condition reflex (Brown & Bellugi, 1964, p. 151."

The proposal to study the acquisition of reading in the same manner, the use of tape recordings of the interactions between the subjects and the teacher, proved impractical. The general classroom noise level interfered with the taping and it became necessary to limit this initial study to general observations and shorthand notes about the interactions of significance. Rather than limiting the observations to two to four subjects after initial visits to the classroom, all the children were observed, their interactions with each other being included as well.

The Classroom

The study was undertaken in the first-grade section of Fairburn School in Victoria, B.C., which was composed of twenty first-graders (12 boys, 8 girls) and, during the afternoon, eight kindergarten children (2 boys, 6 girls). The children were of the appropriate age (5 and 6 years old respectively) and came principally from middle-class homes. The atmosphere in the classroom was friendly and the children seemed eager to respond to

¹ The study was undertaken under the guidance of Dr. Norma I. Mickelson and Dr. Peter O. Evanechko of the Faculty of Education, University of Victoria.

questions. They were accustomed to outside visitors and did not appear to be worried about or overly interested in the presence of an observer.

The programme of instruction was one of integrated language arts combining phonics, sight-word instruction, reading in several texts, and writing.

Data Collection

Observations in the classroom began on February 1st, 1974, and extended to March 20th. Generally, two visits per week were made during that period; however, on several occasions other activities (French lessons, special events, field trips, teacher conferences, etc.) prevented observations of activities connected with the acquisition of reading. Ten sessions of about one-and-a-half hours each provided the material discussed here. Three of the sessions were devoted to work with kindergarten children. As noted above, a written record was kept of interactions between the teacher and pupils. This included blackboard work, oral interactions during reading sessions and exercises in connection with worksheets. These observations were augmented by notations made during oral readings in class and tape recordings of oral readings by three children selected by the teacher and designated as top, middle, and poorest readers among the first-graders. Analysis of the oral readings, both those of the three pupils during special sessions and of groups of children during regular class work, was done on the pattern of the Goodman Miscue Analysis² (Goodman, 1969).

Observed Behaviour Patterns

Analysis of the data collected—records of oral interactions and notations on general reading behaviour—revealed a number of recurring patterns which are listed below with selected examples or explanatory comments.

1. The same wrong answer, or one very similar to it, is frequently given several times in succession by a number of children or even the same child. (The answer looked for by the teacher is usually readily inferred by an adult observer, but apparently not by the children.) The behaviour described in 2 below is a case in point.
2. The child's perceived meaning and his customary names for concrete referents appear to guide his response even after the teacher has rejected an answer several times. The task at hand—say, finding rhyming words or beginning sounds on an exercise sheet—does not intrude itself on the recognition of pictures as “glove” or “beetle” when “mitten” and “bug” are the appropriate responses for that particular exercise.
3. The intended purpose of an exercise—say, learning about letters and sounds—does not necessarily guide the child in working out the answers. On a sheet showing pictures of rhyming words the picture of a cat appeared in four locations. Since “cat” was used to rhyme with “hat” in the first row, a child stated that that meant “cat” was ruled out as a

² In Goodman's view, differences between the expected response and observed response during oral reading yield important insights into the mental processes of the reader. They offer glimpses of reading “from the inside out.” Instead of classifying such differences between the text and the oral rendition of the child as “errors” they are termed “miscues” and used as an analytic tool—with the aid of a taxonomy of miscues—to derive an understanding of the child's thought processes and degree of comprehension of the material read.

rhyming word for the rest of the paper. ("Cat" was, in fact, used only once as a rhyming word in the work sheet.)

4. Statements by other children that "this is the right answer" or demonstrations that "this is the way to do it" appear to be more readily accepted and copied than explanations of an adult—teacher or observer.
Jimmy: How do you spell cauze?
Adult: Do you mean because?
Jimmy: No, CAUZE! (enunciating emphatically)
Other child: Yeah, it's "because," Jimmy.
Jimmy: Oh. O.K., how do you spell because?

5. Explanations provided by the teacher about such concepts as words and sentences frequently fail to elicit a response or sign of recognition. Children appear to deal with language in patterns and semantic units. Classification of the language into discrete grammatical units at the word level has not yet been internalized.

When asked by the teacher to give their favorite word, two of the kindergarten children responded with "happy face" and "swimming lesson." Their responses to the teacher's explanation that those were two words were respectively, "Then I don't know." and a repetition of "swimming lesson." A first-grader faced with a picture and the word "catch" read it as "Are you ready?" in response to the teacher's request to "read the word."

6. "Sounding out" a word does not yet appear to proceed in the manner conceived by the proficient reader. The children frequently appear puzzled about the connection between "the sound" the teacher is demonstrating or talking about and the word they are asked to read. They will cup their lips to imitate the teacher and even produce "the sound," nod "yes" to the question "do you hear the sound?" yet fail to make the transfer from "the sound" to the word.
7. Errors (miscues) during oral reading sessions frequently involve conversion of the text into a language pattern more meaningful in terms of the child's oral language. "Mother" becomes "Mom," "Little Duck" is read as "baby duck," and "No, I have not seen her." is shortened to "No, I haven't."
8. Wrong intonations during oral readings often indicate that the child draws on his knowledge of familiar language patterns or word order to derive meaning from the text. "Father will be back tonight. The children . . ." (In the text, the sentence appeared in the same manner—with "tonight" beginning the next line) becomes "Father will be back. Tonight, the children . . ." The sentence "Then you push a little." created difficulty for a first-grader. He stopped after "push a . . ." despite the fact that "little" is a very familiar word to him. When he was coaxed into reading the sentence his intonation indicated that to him the sentence was not complete. "Little" apparently required a complement.

Discussion

The short duration of the study and the necessity to observe an entire class with little more than the aid of a pencil and paper made it impossible to draw definitive conclusions. However, even these limited observations

confirmed that there are indeed parallels between the processes of language acquisition and learning to read. Brown and Bellugi (1964) found that some of the ungrammatical utterances produced by the infants they observed yielded particularly interesting data for analysis since the "inductive operations of the child's mind are externalized in such a creation." Similarly, reading errors and the children's overall reactions to instruction in the classroom at times externalized their thinking. It revealed that from a very early stage, beginning readers, like infants learning to speak, draw upon the totality of their environment and experience to "make sense" of language: in the classroom, children look to the teacher and to other children to confirm the correctness of their efforts with written language; they examine pictures in the text, and, above all, they draw on their knowledge of their own patterns of oral language.

Where for the infant, contentives (nouns, verbs, and adjectives) dominate early speech, and language patterns (word order, noun phrases, etc.) appear to guide the further development of language, so the beginning reader seems to rely on concrete referents and his own language patterns to deal with written language (see examples under 2, 5, 7, & 8 in the foregoing section.)

Oral practice rather than reliance on explanations or abstract concepts appears to guide the learners in most cases. The infant practises speech on the basis of the oral language he hears around him. He receives feedback on his efforts and gradually approximates the adult models' speech more closely. He himself abstracts the rules, tests them orally, (Child: Daddy goed out. Mother: That's right, Daddy went out.) and eventually gains proficiency in their use; yet the rules themselves are never explicitly stated. In class, beginning readers repeat the teacher's examples, receive feedback on their own performance and that of other children in the class. Their ability to perform correctly improves gradually even though they cannot accurately verbalize the explanation the teacher has repeated many times. During exercises involving word endings, most children performed correctly but were unable to state reasons for their correct responses. Examples collected several weeks apart revealed that proficiency in the use of apostrophe s to indicate possession had improved, yet the children continued to use a tautology to explain their choice: "There is an apostrophe s on the end because it's Ann's doll." None repeated the teacher's explanation, "The doll belongs to Ann."

During phonics lessons, children perform correctly in a high proportion of cases but display a certain rigidity of response. Often, the more proficient readers will provide the initial examples and others follow suit. However, at times, responses quite unlike those desired are given and the teacher's question, "Do you hear the sound?" brings out the fact that the connection between "the sound" and the word is not perceived by a child who can nevertheless read.

Since the child is unclear about such concepts as sounds, words, sentences, and context, it becomes apparent that he largely abstracts and integrates the rules of reading on the basis of his own cue system aided by the oral practice in class. Much like the child learning to speak, the beginning reader inductively arrives at latent structures in the language (in this

instance written language). In this process, the explanations offered by the teacher may take a role akin to that of the mother's expansion of the child's utterances during the "imitation-with-expansion" phase of the process of language acquisition. As the child is practicing his new skills in class, the teacher acts as "communication check" to confirm, expand on, or correct his efforts.

On the basis of research now under way, further analogies between the processes of language acquisition and reading acquisition are beginning to emerge. In addition to Brown and Bellugi's study, some facets of Lenneberg's (1969) biological theory of the development of language should prove useful in formulating a more rigorous theoretical basis for the analysis of data collected during classroom observations. His concepts of gradual development, the actualization of latent structures into realized structures, the function of the social environment as a triggering mechanism (rather than a causal agent), and the process of assimilation, could readily be applied to a model of reading acquisition.

Conclusion

Analysis of data collected during observations in a first-grade classroom appears to confirm that there is a similarity between the processes of learning a language and learning to read, though the input provided for learners differs. The beginning reader is guided by explanations regarding phoneme-grapheme correspondences and the rules of syntax and grammar. The infant learning to speak, acquires proficiency without the benefit of explanations or stated rules. Both have the opportunity to practice orally and both receive feedback on their efforts. The finding that despite the differences in the learning environments, beginning readers nevertheless appear to use strategies similar to those of the beginning oral language user, raises the question to what extent the child in the classroom needs and uses the rules presented to him during the initial efforts to penetrate written language.

If reading is defined as the process of deriving meaning from the printed word, then an untimely introduction of rules and abstract explanations may actually retard the process of learning to read. While the child is searching for meaning, we, the teachers, are fixating on "sounds", letters, and words. As Frank Smith (1974) put it, "If we tried to teach children to speak the way we teach them to read, they might never learn." Assuming the necessity to "know the rules" before being able to perform, seems to be putting the cart before the horse. The external input presented by the teacher serves as a model and provides that all-important feedback on the child's performance; however, it would appear that it is not that the child learns to read by sounding out words, but rather that he learns to sound out words once he has learned to read. He knows the right answers before he knows why he is right. He learns the language of reading before he learns the rules.

References

- Brown, R., & Bellugi, U. Three processes in the child's acquisition of syntax. *Harvard Educational Review*, 1964, 34, 133-151.
- Burke, C. L., & Goodman, K. S. When a child reads: A psycholinguistic analysis. *Elementary English*, 1970, 47 (1).
- Goodman, K. S. A linguistic study of cues and miscues in reading. *Elementary English*, 1965, 42.
- Goodman, K. S. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly*, 1969 (Fall).
- Goodman, Y., & Burke, C. Do they read what they speak? *Grade Teacher*, 1969, 26(7).
- Lenneberg, E. H. *Biological foundations of language*. New York: John Wiley and Sons, Inc., 1967.
- Smith, E. B., Goodman, K. S., & Meredith, R. *Language and thinking in the elementary school*. New York: Holt, Rinehart & Winston, 1970.
- Smith, F. *Understanding reading*. New York: Holt Rinehart & Winston, 1971.
- Smith, F. *Psycholinguistics and reading*. New York: Holt Rinehart & Winston, 1973.
- Smith, F. Address given at the University of Victoria, Victoria, Canada, Spring, 1974.

Editor's Note:

It has not been a general policy of AJER to publish bibliographies. However, this one, developed over the last several years by Dr. Blackall, represents an area which we think will be of wide interest to our readers.

BETTINA M. BLACKALL

The University of Alberta

The Works of Jean Piaget Published in English

Piaget, J. *The language and thought of the child.*

Translated by Marjorie Gabain. London: Routledge & Kegan Paul, 1959. Original French edition, 1923; First published in English, 1926.

London: Routledge & Kegan Paul, 1926, 1948.

New York: Harcourt, Brace, 1928.

London: Routledge & Kegan Paul, 1952.

New York: World Books, 1955.

New York: Meridian, 1957.

London: Routledge & Kegan Paul, 1959. (pb)

New York: Philosophical Library, 1961.

New York: Humanities Press, 1962. (Reproduction of the 1929 ed.)

Piaget, J. *Judgment and reasoning in the child.*

Translated by Marjorie Worden. New York: Littlefield 1968. Original French edition, 1924; First published in English, 1928.

London: Routledge & Kegan Paul, 1928.

New York: Harcourt, Brace, 1928, 1938.

London: Routledge & Kegan Paul, 1951.

New York: Humanities Press, 1962. (Reproduction of the 1928 ed.)

New York: Littlefield, 1968, (pb)

Piaget, J. *The child's conception of the world.*

Translated by Joan & Andrew Tomlinson. New York: Littlefield, 1969. Original French edition, 1926; First published in English, 1929.

London: Routledge & Kegan Paul, 1929.

New York: Harcourt, Brace, 1929.

London: Routledge & Kegan Paul, 1951.

New York: Humanities Press, 1960. (Reproduction of the 1929 ed.)

Totowa, N. J.: Littlefield, Adams, Patterson, 1960.

New York: Littlefield, 1969. (pb)

London: Paladin, 1973. (pb)

Piaget, J. *The child's conception of physical causality.*

Translated by Majorie Gabain. New York: Littlefield, 1965. Original French edition, 1927; First published in English, 1930.

London: Routledge & Kegan Paul, 1930.

New York: Harcourt, Brace, 1930.

Totowa, N. J.: Littlefield, Adams, Patterson, 1960.

New York: Littlefield, 1965. (pb)

New York: Humanities Press, 1966. (Reproduction of the 1930 ed.)

Piaget, J. Children's philosophies. In C. Murchison (Ed.), *A handbook of child psychology*. Barre: Clark University Press, 1931. Pp. 337-391.

Piaget, J. Retrospective and prospective: Analysis in child psychology. *British Journal of Educational Psychology*, 1931, 1, 130-139.

Piaget, J. *The moral judgment of the child.*

Translated by Marjorie Gabain. New York: Macmillan, 1955. Original French edition, 1932; First published in English, 1932.

London: Routledge & Kegan Paul, 1932.

New York: Harcourt, Brace 1932.

New York: Macmillan, 1955.

Piaget, J. *Social evolution and the new education.* (Vol. 8) London: New Education Fellowship, 1932.

Piaget, J. Principal factors determining intellectual evolution from childhood to adult life. In *Factors determining human behavior*. Harvard Tercentenary Publication, I. Cambridge, Mass.: Harvard University Press, 1937. Pp. 32-48.

Piaget, J. Intellectual evolution. In *Science and man*. Totowa, N. J.: Littlefield, Adams, Patterson, 1942. Pp. 409-422.

Piaget, J., & Inhelder, B. Diagnosis of mental operations and theory of intelligence. *American Journal of Mental Deficiency*, 1947, 51, 401-406.

Piaget, J. *The psychology of intelligence.*

Translated by M. Piercy & D. E. Berleyne. New York: Littlefield, 1968. Original French edition, 1947; First published in English, 1950.

London: Routledge & Kegan Paul, 1950.

New York: Harcourt, Brace World, 1950.

Totowa, N. J.: Littlefield, Adams, Patterson, 1960.

New York: Littlefield, 1968. (pb)

Piaget, J., & Weil, A. M. The development in children of the idea of the homeland and of relations with other countries. *International Social Science Bulletin*, 1951, 3, 561-578.

Piaget, J. *Play, dreams and imitation in childhood*.

Translated by C. Gattegno & F. M. Hodgson. New York: Norton, 1962. Original French edition, 1945; First published in English, 1951.

London: Routledge & Kegan Paul, 1951.

London: Heinemann, 1951.

New York: Norton, 1952.

New York: Norton, 1962. (bp)

Piaget, J. The right to education in the modern world. In UNESCO. *Freedom and culture*. New York: Columbia University Press, 1951. Pp. 67-116.

Piaget, J. *The child's conception of number*.

Translated by C. Gattegno & F. M. Hodgson. New York: Norton, 1965. Original French edition, 1941; First published in English, 1952.

London: Routledge & Kegan Paul, 1952, 1961.

New York: Humanities Press, 1952.

New York: Humanities Press, 1962, 1964.

London: Routledge & Kegan Paul, 1965. (pb)

New York: Norton, 1965. (bp)

Piaget, J. Jean Piaget. In E. G. Boring *et al.* (Eds.), *A history of psychology in autobiography*. Vol. IV Barre: Clark University Press, 1952. Pp. 237-256.

Piaget, J. *The origin of intelligence in children*.

Translated by Margaret Cooke. New York: Norton, 1963. Original French edition, 1936; First published in English, 1952.

New York: International Universities Press, 1952, 1966.

London: Routledge & Kegan Paul, 1953.

New York: Norton, 1963. (pb).

Piaget, J. Genetic psychology and epistemology. *Diogenes*, 1953, 1, 49-63.

Piaget, J. How children form mathematical concepts. *Scientific American*, 1953, 189(20), 74-78.

Piaget, J. *Logic and psychology*.

Translated by W. Mays & F. Whitehead. (Based on lectures delivered at the University of Manchester, 1952.) New York: Basic Books, 1957. Original French edition, 1953; First published in English, 1953.

Manchester: The University Press, 1953.

New York: Basic Books, 1957.

- Piaget, J. *The construction of reality in the child*.
Translated by Margaret Cooke. New York: Ballantine, 1971. Original
French edition, 1937; First published in English, 1954.
New York: Basic Books, 1954.
London: Routledge & Kegan Paul, 1955.
New York: Ballantine, 1971. (pb)
- Piaget, J. Language and thought from the genetic point of view. *Acta Psychologica*, 1954, 10, 51-60.
- Piaget, J. Problems of consciousness in child psychology: Developmental changes in awareness.
Translated by E. Meyer. In H. A. Abramson (Ed.), *Problems in consciousness*. Transactions of the Fourth Conference, 1953, Princeton, N. J. New York: J. Macey Foundation, 1954. Pp. 136-177.
- Piaget, J. The development of time concepts in the child. In R. H. Hoch, & J. Zubin (Eds.), *Psychopathology of childhood*. New York: Grune & Stratton, 1955. Pp. 34-44.
- Piaget, J. Perceptual and cognitive (or operational) structures in the development of the concept of space in the child. *Acta Psychologica*, 1955, 11, 41-46.
- Piaget, J., & Inhelder, B. *The child's conception of space*.
Translated by F. J. Langton & J. L. Lunzer. New York: Norton, 1967.
Original French edition, 1948; First published in English, 1956.
London: Routledge & Kegan Paul, 1956.
New York: Humanities Press, 1963.
New York: Norton, 1967. (pb)
- Piaget, J. Some impressions of a visit to Soviet psychologists. *American Psychologist*, 1956, 11, 343-345.
- Piaget, J. Some impressions of a visit to Soviet psychologists. *Acta Psychologica*, 1956, 12, 216-219.
- Piaget, J. The child and modern physics. *Scientific American*, 1957, 196(Mar.), 46-51.
- Inhelder, B. & Piaget, J. *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*.
Translated by Anne Parsons & S. Milgram. New York: Basic Books, 1958. Original French edition, 1955; First published in English, 1958.
- Piaget, J., Vinh-Bang, & Matalon, B. Note on the law of the temporal maximum of some optico-geometric illusions. *American Journal of Psychology*, 1958, 71, 277-282.

- Piaget, J. Peering into the mind of a child. *The UNESCO Courier*, 1959, 12, 4-7.
- Piaget, J., Inhelder, B., & Smeminska, A. *The child's conception of geometry*.
Translated by E. A. Lunzer. New York: Basic Books, 1960. Original French edition, 1948; First published in English, 1960.
London: Routledge & Kegan Paul, 1960.
New York: Basic Books, 1960.
- Piaget, J. Individual and collective problems in the study of thinking. *Annals of the New York Academy of Sciences*, 1960, 91, 22-37.
- Piaget, J. The genetic approach to the psychology of thought. *Journal of Educational Psychology*, 1961, 52, 275-281.
- Piaget, J. The language and thought of the child. In *Classics in psychology*. New York: Thorno Shipley, 1961. Pp. 994-1031.
- Piaget, J. *Comments on Vygotsky's critical remarks concerning "The language and thought of the child" and "Judgment and reasoning in the child"*. Cambridge, Mass.: M.I.T. Press, 1962.
- Piaget, J. Three lectures. (The stages of the intellectual development of the child; The relation of affectivity to intelligence in the mental development of the child; Will and action.) *Bulletin of the Menninger Clinic*, 1962, 26, 120-145.
- Piaget, J. The attainment of variance and reversible operations in developmental thinking. *Scientific American*, 1963, 30, 283-299.
- Piaget, J. Cognitive development in children: The Piaget papers. (Development and learning; The development of mental imagery; Mother structures and the notion of number; Relations between the notions of time and speed in children.) In R. E. Ripple & V. N. Rockcastle (Eds.), *Piaget rediscovered*. A Report of the Conference on Cognitive Studies and Curriculum Development, March 1964. Ithaca: Cornell University Press, 1964.
- Piaget, J. Development and learning. *Journal of Research in Science Teaching*, 1964, 2, 176-186.
- Inhelder, B., & Piaget, J. *The early growth of logic in the child: Classification and seriation*.
Translated by E. A. Lunzer & D. Papert. New York: Norton, 1969. Original French edition, 1959; First published in English, 1964.
London: Routledge & Kegan Paul, 1964.
New York: Harper & Row, 1964.
New York: Norton, 1969. (pb)

- Piaget, J. Psychology and philosophy.
Translated by R. Howard. In Wohlman & Nagel (Eds.), *Scientific psychology, Principles and approach*. New York: Basic Books, 1965. Pp. 28-43.
- Piaget, J. Foreward to M. Almy *et al.*, Young children's thinking: Studies of some aspects of Piaget's theory. New York: Teacher's College Press, Columbia University, 1966.
- Beth, E. W., & Piaget, J. *Mathematical epistemology and psychology*. New York: Gordon & Breach, 1966. Original French edition, 1961; First published in English, 1966.
- Piaget, J. Response to Brian Sutton-Smith. *Psychological Review*, 1966, 73, 111-112.
- Piaget, J. Time perception in children.
Translated by B. Montgomery. In J. P. Frazer (Ed.), *The voices of time*. New York: Braziller, 1966. Pp. 202-216.
- Piaget, J. Cognitions and conservations. *Contemporary Psychology*, 1967, 12, 532-533.
- Piaget, J. *Six psychology studies*. Edited by D. Elkin;
Translated by Anita Tenzer. New York: Vintage Books, 1969.
Original French edition, 1964; First published in English, 1967.
New York: Random House, 1967.
London: The University of London Press, 1969.
New York: Vintage Books, 1969. (pb)
- Piaget, J. *Studies in cognitive growth* (Bruner *et al.*):
Review. *Contemporary Psychology*, 1967, 12, 532-533.
- Piaget, J., *et al.* *Experimental psychology: Its scope and method. Vol. I. History and method*. New York: Basic Books, 1968.
- Piaget, J. *On the development of memory and identity*. (Heinz Werner Lecture Series, Clarke University, Worcester, Mass. Vol. 2.) New York: Columbia University Press, 1968.
- Piaget, J. Quantification, conservation and nativism. *Science*, 1968, 162, 976-979.
- Piaget, J. *The child's conception of movement and speed*.
Translated by G. E. T. Holloway & M. J. Mackenzie. New York: Ballantine, 1971. Original French edition, 1946; First published in English, 1969.
New York: Basic Books, 1969.
London: Routledge & Kegan Paul, 1970.
New York: Ballantine, 1971. (pb)

Piaget, J. *The child's conception of time.*

Translated by A. J. Pomerans. New York: Ballantine, 1971. Original French edition, 1927; First published in English, 1969.

London: Routledge & Kegan Paul, 1969.

New York: Basic Books, 1970.

New York: Ballantine, 1971. (pb)

Piaget, J. *Mechanisms of perception.*

Translated by G. N. Seagram. London: Routledge & Kegan Paul, 1969. Original French edition, 1961; First published in English, 1969.

London: Routledge & Kegan Paul, 1969.

New York: Basic Books, 1969.

Piaget, J., & Inhelder, B. The gaps in empiricism. In A. Koestler (Ed.), *Beyond reductionism: The Alpbach Symposium*, 1968. London: Hutchinson, 1969.

Piaget, J. & Inhelder, B. *The psychology of the child.*

Translated by Helen Weaver. New York: Basic Books, 1969. Original French edition, 1966; First published in English, 1969.

New York: Basic Books, 1969.

New York: Harper Torchbooks, 1969. (pb)

London: Routledge & Kegan Paul, 1969.

Piaget, J. *Genetic epistemology.*

Translated by Eleanor Duckworth. Woodbridge Lectures, 1968. New York: Norton, 1971. First published in English, 1970.

New York: Columbia University Press, 1970.

New York: Norton, 1971. (pb)

Piaget, J. *Genetic psychology.*

Translated by Eleanor Duckworth. Ithaca: Cornell University Press, 1970.

Piaget, J. Piaget's theory. In P. H. Mussen (Ed.), *Carmichael's Manual of child psychology*, Vol. I. New York: Wiley, 1970. Pp. 703-732.

Piaget, J. *Science of education and the psychology of the child.* New York: Viking, 1971. First published in English, 1970.

New York: Orion Press, 1970.

New York: Viking, 1971 (pb)

Piaget, J. *Structuralism.*

Translated by Channah Maschler. New York: Harper Torchbooks, 1971. Original French edition, 1968; First published in English, 1970.

New York: Basic Books, 1970.

New York: Harper Torchbooks, 1971. (pb)

Piaget, J. *Biology and knowledge: An essay on the relations between*

- organic regulations and cognitive processes*. Edinburgh: Edinburgh University Press, 1972. First published in English, 1971.
Chicago: University of Chicago Press, 1971.
Edinburgh: Edinburgh University Press, 1972.
- Piaget, J. *et al. Mental imagery in the child: A study of the development of imaginal representation*.
Translated by P. A. Chilton. London: Routledge & Kegan Paul, 1971.
Original French edition, 1966; First published in English, 1971.
- Piaget, J. The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer (Eds.), *Measurement and Piaget*, Proceedings of the CTB/McGraw Hill Conference on Ordinal Scales of Cognitive Development. New York: McGraw-Hall, 1971. Pp. 1-11.
- Piaget, J. Comments on mathematical education. Paper presented at the International Conference for Mathematics Education, Exeter, 1972.
- Piaget, J. *Insights and illusions of philosophy*.
Translated by W. Mays. London: Routledge & Kegan Paul, 1972.
First published in English, 1972.
- Piaget, J. Intellectual evolution from adolescence to adulthood. *Human Development*, 1972, 15, 1-12.
- Piaget, J. Play and development. In M. W. Piers (Ed.), *Play and development: A symposium*. New York: Norton, 1972.
- Piaget, J. *The principles of genetic epistemology*.
Translated by Mays Wolfe. New York: Basic Books, 1972. First published in English, 1972.
- Piaget, J. Problems of equilibration. In C. F. Nodine, J. M. Gallagher, & R. H. Humphreys (Eds.), *Piaget and Inhelder: On equilibration*. Philadelphia: The Jean Piaget Society, 1972. Pp. 1-20.
- Piaget, J. *Psychology and epistemology: Towards a theory of knowledge*.
Translated by Arnold Rosin. New York: Viking, 1972. First published in English, 1972.
- Piaget, J. The role of imitation in the development of representational thought. *International Journal of Mental Health*, 1972-3, 1, 67-74.
- Piaget, J. *The child and reality: Problems of genetic psychology*.
Translated by Arnold Rosin. New York: Grossman, 1973. Original French edition, 1972; First published in English, 1973.
- Piaget, J., & Inhelder, B. *Memory and intelligence*.
Translated by Arnold J. Pomerans. London: Routledge & Kegan Paul, 1973. Original French edition, 1968; First published in English, 1973.



ISSN 0002-4805

The Alberta Journal of Educational Research

Volume XXI, No. 2

June, 1975

CONTENTS

Educational Research: Concepts and Models	71
<i>G. R. Eastwood</i>	
A Study of a Public School System, Hessen, West Germany	84
<i>D. Friesen</i>	
Status Consensus as it Relates to Team Teacher Satisfaction	104
<i>M. S. Arikado</i>	
Student Perceptions of Teacher Pupil Control Behavior and Student Attitudes Toward High School	110
<i>W. Pritchett and D. J. Willower</i>	
Patterns of Learning Projects Among Professional Men	116
<i>C. A. McCatty</i>	
Reading or Listening: A Comparison of Techniques of Instruction in Elementary Social Studies	130
<i>Joan C. Breiter</i>	
List of Books Received	141

FACULTY OF EDUCATION
The University of Alberta

G. R. EASTWOOD

Simon Fraser University

Educational Research: Concepts and Models

The author of this essay notes and accepts that educational practice has not benefited sufficiently from research. He argues that educational research conceived as operational research has a refining and reformulating rather than a discovery function. A conceptual model designed to present research as a systematically formulated integration of theoretical—empirical—practitioner components is developed and explained. Use is made of an error eliminating procedure for the progressive development of educational ideas and theories. Implicit in the argument is the thesis that educational (and teaching) innovations do not arise from ideas that are inductively generated by research or practice—the function of these two activities is to structure and refine ideas prior to and during implementation. The ideas are the result of informed guesses or conjectures. Implicit too is a plea that there be adequate attempts to refine the ideas before large scale implementation is undertaken. The essay concludes with a brief statement about an appropriate form of organization for research units. (Dr. Eastwood is Professor in the Faculty of Education, Simon Fraser University, Burnaby.).

There is widespread belief, and probably ample supporting evidence, that educational practices and developments have not benefited significantly from the input of resources that, particularly over the past ten or fifteen years, have been assigned to educational research. Some argue that the resources committed both in terms of needs and in terms of comparison with other fields although large are inadequate. Others take the view that the problems are not resolvable through research and that the resources could be better employed in other ways. Some take consolation from the belief that educational research (in common with social science research generally) is in its infancy and that it is unreasonable to expect positive results at this stage. There are others who would simply argue that the research has been inadequately conceived, designed and executed. It is the case that a large proportion of the experimental designs used have been borrowed from other fields and used with little or no modification. Classical examples are to be found in the use of complex statistical designs developed for agricultural research (e.g. randomized block designs). All of

these reactions may have some validity but debating their causal relationship to the present state is a pointless activity and not pertinent to the purpose here.

Until quite recently educational philosophers have largely ignored educational research as a field of philosophical activity. In view of the very rapid and very significant development of the philosophy of science (with its several facets) during recent decades, it does seem remiss that educators and particularly educational philosophers have not turned their attention to educational research and its conceptual and methodological problems. Some work has been done recently. The publication of *Philosophical Redirection of Education Research* (1972) contributed very little. It has been reviewed elsewhere (*Educational Theory*, 1973) and need not be the subject of further comment here. *Philosophy of Educational Research* (Broudy, Ennis, & Krimmerman, 1973) is a significant contribution and, additionally, a firm indication that philosophers have begun to examine the problems. The essays in Chapter 2 of that volume present points of view regarding "whether we should think of educational research as laying the foundation of a science in its own right, or should it be considered a search for ways of applying the findings of psychology, sociology, anthropology and other behavioural sciences?" (Broudy et al., 1973, p. 65). The main thrust of the chapter is therefore directed to the distinction between educational research as basic research and as applied or operational research. The consensus of these essays tends to be on the latter.

Additionally and more importantly there seems to be agreement that educational progress is dependent more upon what teachers do and how they do it than on any other factor. This has, as one contributor says, implications for the "theoretical level at which prospective teachers are trained" (Kerlinger, 1973, pp. 102-107).

In this article, it is argued that educational research is (a) primarily concerned with operational research, (b) a critical refining and reformulating rather than a discovering activity, and (c) a systematically organized joint philosopher-researcher-practitioner enterprise.

That we act or behave in any context in a manner consistent with the way in which we conceptualize that context has been well established. From this it follows that giving direction to an enterprise or providing redirection is not a matter of formulating and establishing rules but of creating a conceptual framework—a conceptual system that causes those within it to act appropriately. That is what this paper attempts to do; the reader must estimate the degree of success achieved.

Discoveries, innovations, developments or knowledge growth generally are seldom the specific result of research. We do not "do" research, discover something and then use or apply it. The belief that research or theory can be applied in practice or used to guide practice according to a linear model can be called the educators' fallacy. Neither the generation of new ideas nor the formulation of tentative solutions to problems is, in any field of inquiry, an inductive process. Nor do they occur as a result of deductive reasoning. They are not the result of logical processes in the usual sense of the term. They are for the most part what, in K. R. Popper's terms, can be called *conjectures*. They occur as alternatives to existing procedures with which there is either general dissatisfaction, or for which experience has

shown or tends to suggest they may be more appropriate. Generally they are imaginative conjectures which are formulated by individuals who are knowledgeable in the field and who have a “hunch” that some other way might be preferable and more effective. They tend to be conjectures in a weaker sense than their counterparts in scientific fields but there seems to be sufficient similarity for them to be called conjectures; e.g. the discovery method idea, the team-teaching idea and a host of others past and present. In advancing these, their advocates simply *conjecture* that they are or are likely to be more effective than those they displace. It is stressed here that they are conjectures and that they are not theories. At the same time it must be noted that theories begin as conjectures and these innovative educational ideas may be the seeds from which theories grow.

The second assumption or premise concerns the relationship between the knowledge available and the knowledge actually used in educational contexts. For example, in spite of the reservations we hold about its universal validity, there is available in psychology a great deal more knowledge about how and why people learn and the conditions under which learning is likely to be maximum than teachers are able to incorporate in educational situations. This is true for a wide range of areas in psychology, in sociology, in curriculum studies and others. Utilization of the facts known and the processes developed is largely a matter of chance and dependent upon the creative ability of individuals—individuals, be it noted, who are for the most part neither as carefully selected nor as highly qualified as they need to be to effect transfer to practice without aid; individuals who are required to work under constant day-to-day pressures in a practical endeavour and who are accorded little time for reflection before they have to act.

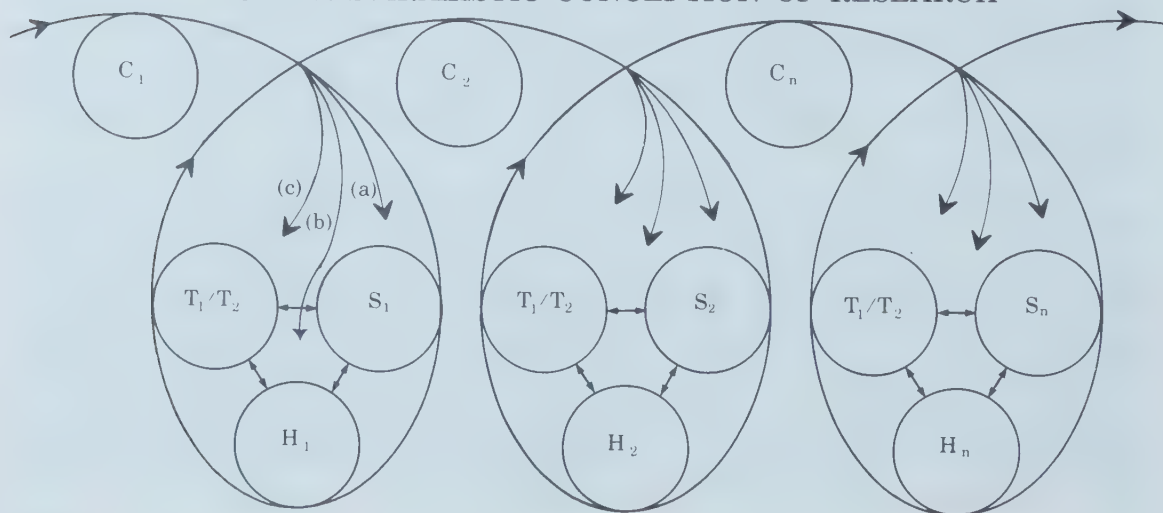
This is an extremely important factor which points out the need for the development of comprehensive factually oriented theories of teaching. Educators have too long laboured under the misapprehension that it is the function of theory to guide practice. By so doing they have deprived the teacher of the benefits from theory's role in accounting for the known facts and maximizing the likelihood of their being coherently assembled and systematically used. We do have available a great deal of information regarding the function of practice, of reinforcement, of motivating techniques, of curiosity. We also know a great deal about sociological factors that influence cognitive development and we have available information regarding ways of organizing and presenting subject matter content. What we do not know or have not made available to teachers are systematic ways of putting all these “knowings” and “understandings” together in a teaching situation.

There are fields in which it is appropriate for the researcher to operate independently from the practitioner. In these fields the process or product is developed or discovered and then made available to practitioners. But when application is to be made to human beings, the situation becomes more complex partly because all do not respond in the same way and partly because appropriate, varied and flexible modes of application have to be developed. Developing these is part, perhaps the greater part, of educational research. Discovery and utilization (research and practice) are two aspects of a single endeavour.

No one would wish to suggest or be accused of suggesting that a technological research pattern is an adequate or appropriate model for educational research, but there is one aspect that is worthy of consideration. Human learning and human development are considerably more complex processes than the development of a new machine. But both are practical activities dependent upon the application and utilization of basic research and both necessitate research in ways of applying. Consider the painstaking care with which a machine (e.g. a new and innovative design) is developed from the point of conception of the idea through careful analysis of the idea, through scale models of parts and of the whole, through prototype after prototype until the final product is produced. Then consider the extensive series of tests and modifications that both precedes and accompanies introduction to service. Then acknowledge that in spite of it all, faults and weaknesses appear—sometimes very serious ones. Contrast this long expensive process with the simplistic assumption that an educational innovation can be conceived and introduced (often on a very large scale) without the safeguards of carefully controlled trials of a series of prototypes. And yet we all know and admit freely that the processes of education are much more complex than those of technology. Out of this arises the theory that educational research is, or ought to be, an integrated activity involving the continuous interplay of the conceptual, empirical and practical dimensions. That, of course, means not only the integrated efforts of philosophers, social scientists, educational researchers and practitioners but the blurring of the distinctions and divisions between and among them.

Figure 1 is an attempt to represent diagrammatically a research and educational development model which is designed to conceptualize educational research, educational philosophy and educational practice as complementary processes. At the same time it is designed to facilitate a conception of educational development as having inseparable philosophical and empirical aspects—as being conceptual and empirical without separation and division of the two. No doubt it could have been formulated in terms of the disciplines and personnel to be used. It has been structured here in terms of the kinds of activities involved, but whether or not the right selection has been made is uncertain.

FIGURE 1. SPIRALISTIC CONCEPTION OF RESEARCH



The elements of Figure 1. need considerable and careful description and explanation. First, the symbols will be identified and given very brief descriptions.

- C_1, C_2, \dots, C_n = Sequence or set of conjectures which may or may not be cumulative
 $S_1 S_2 \dots S_n$ = Systematic theoretical formulations derived from C
 $H_1 H_2 H_n$ = Triable or testable hypotheses exemplifying some aspects of or derivative from S
 $T_1 T_2 \dots T_n$ = Testing situations—empirically based but more often experiential trials than experimental tests.

At this point the model and the elements therein appear as a simplistic model, for the sake of a model, of a naive conception of scientific method. That it is not must be shown, and to begin, this will be done by considering the diagrammatic representation as a whole—or at least one cycle of it. Successive cycles are included only to show that this is not a final solution context. People and societies change and solutions satisfactory today are not likely to remain satisfactory tomorrow. Some of its characteristics have occurred accidentally and spontaneously (as ideas or conjectures generally do) and by so doing have provided an analogy that seems to aptly illustrate the concept and to enable us to see it as an organic whole.

Conjecture C_1 is an idea—a complex of many undefined and undifferentiated elements. It is external to but on a trajectory that leads into the loop containing S_1, H_1 and T_1 , which are interrelated and interconnected. The ideas or complex of ideas embodied in C are diffuse, undifferentiated and only vaguely specified. Any one or any group may become the germinal constituent that set in motion the S-H-T process. The analogy that is suggested is that of biological fertilization and growth. It is a good analogy because it implies a concept of S-H-T as an organic developing whole and not as a series of stages. Using such an analogy leads to conceptualization of C as seminal and containing a large number of elements only one of which may initiate the research and development process. Past observations of the exemplifications in practice of conjectures lead to the conclusion that we have not had a great deal more control over which element in C initiated development than we have over which sperm initiates cellular growth. For example there does seem to be an almost infinite variety of exemplifications of “discovery method” learning, all supposedly derived from the same idea or principle. What is done or created in the name of that principle depends upon the conception of it in the mind of the developer.

At this point the analogy breaks down because while it is not likely that under natural conditions any control over which sperm initiates development can be achieved this is not the case with the seminal idea C. If educational research is to be validly exemplified in practice by making an impact on practice then the first stage is that of conceptual analysis and re-synthesis of the conjecture that initiates it. This process is fundamentally philosophical but dependent upon the matrix of empirical concepts and empirical data. This is the process that determines theoretically what will count as an exemplification or application, and what will not count. It is important to note that the specification of those forms that do not count as exemplifications is the most important aspect and the one that has been

most neglected. This point must be stressed. The identification of forms of activity that are not consistent with the idea is of major importance because (a) it provides a safeguard against subjectivism, (b) it provides a context for critical argument, and (c) it leads ultimately to error elimination. Educational ideas (innovations) that fail do so because they are generally so broadly and vaguely formulated that they do not preclude the occurrence of invalid exemplifications.

The research cycle proper—the S-H-T cycle within the loop—is represented in the figure in such a way as to indicate an interconnectedness rather than a sequential order leading from S through H to T and thence into a new cycle. The refined versions of C that are the products of the process of conceptual analysis, clarification and synthesis may lead, as the diverging arrows suggest, to any point in the S-H-T cell. Arrow (a) refers to the simple state wherein a careful (though perhaps very demanding) analysis of C leads to the delineation of a set of elements which are systematically structured and ordered to constitute S_1 , or more likely a set $S_{11}, S_{12} \dots S_{1n}$. It is not likely that in a field as complex as this that one and only one S will occur. When this is the case and an S_1 set can be specified then the process of deriving H_1 or an $H_{11}, H_{12} \dots H_{1n}$ set can be realized. It is a matter of deductive logic but not in the sense of it necessarily being the case that a single H_1 is deducible from a particular S_1 . Nor is it the case that $H_{11}, H_{12} \dots H_{1n}$ are empirically testable in the strict sense of an experimental test. There may be H that are not empirically testable at all and it may be that such H have to be ignored. However the more likely outcomes are H that are testable either experimentally or experientially, or both. This is why in the description of the diagram of the model, T was designated as “testing situations”, not testable hypotheses.

To some degree, though undoubtedly not to a sufficient degree, the C, S and T aspects of the model are already in evidence. They are not sufficiently purposefully established and organized particularly in the tremendously important educational innovative development area, but they are there. We do to some degree conjecture, systematize and hypothesize. It is the testing situation aspect that most needs consideration and it must be separated into at least two phases or stages. This is important if the field of educational practice is, on the one hand, to benefit from research and on the other, to provide the context for refinement of the conclusions from basic research, i.e. become an integral part of the research process. Throughout the above, basic research, including research in fields such as psychology and the social sciences, has been assumed—it is part of the context out of which conjectures arise.

Experiential Triability

Two forms or categories of testing situations are necessary and normally, but perhaps not always, they run sequentially. Whether the initial conjecture C and the subsequent systematization S leading to some hypothetical proposition H is in the context of a specific research topic or in the broader context of a methodological procedure there is need for initial trial in an idealized situation—or one as ideal as possible. It may be and usually is the case that strictly controlled experimental tests cannot be carried out in educational research. There are too many variables that

can't be controlled and too many relations between them that probably can't be identified. But it does not follow that rigorous testing of a well formulated idea cannot be undertaken under conditions as idealized as practical situations will allow. A new teaching method or an educational policy like a new type of machine must be useable and effective under ordinary service conditions—by ordinary teachers in ordinary teaching environments. But those teachers in those environments can no more carry out the initial trials than can an ordinary airline pilot discharge the specialized duties of the test pilot. Educators and educational researchers have failed to see this and as a result have attempted to move from conjecture to practice without the aid of the intervening stages. Providing those stages necessarily implies conceptualizing the research and development process as a complex whole and making positive provision for the necessary stages. The moves from C through S_1 to H_1 are largely but perhaps not wholly logico-philosophical, but T_1 are predominately empirical though admittedly not experimental in the strict sense.

The T phase operates at two levels T^1 and T^2 . The first, T^1 , and ideally speaking it is temporarily foremost, is the "as near as possible" ideal situation in which the hypothesized process or procedure is tried out under conditions as carefully controlled as possible. These are not likely to be truly laboratory conditions but they are carefully designed, staffed and organized environments that maximize to a very high degree the probability of the process or procedure being tested being consistent with the S from which the T has been derived. They may be either specially selected or specially designed environments (classrooms, schools, workrooms etc.) into which specially selected personnel are directed. They are designed to produce situations that coincide closely with the requirements of the H that is under consideration and hence to maximize the probability of confirmation. But if designed or allowed to do only that they will contribute little. The whole purpose of the exercise is to refine H, and hence S and C, and that purpose will not be attained if H is made or shown to work. Some have argued that in educational research it is not possible to formulate a falsifiable hypothesis. This is not the case. The history of educational practice is a chronology of innovations and renovations introduced with enthusiasm and confidence, maintained for periods of varying lengths and eventually either deliberately rejected or discontinued because they have fallen into disfavour. That is falsification but falsification after full scale utilization which is akin to the withdrawal from service of the inadequately tested airliner after a series of fatal accidents. (Generally the airliner is returned to service after modification—the educational counterpart tends not to be modified but to be replaced by another equally untested area or innovation.) Unfortunately the fatal accidents of educational innovations are neither as easily detected nor taken as seriously as are airline accidents.

If as a result of T^1 (a single test or trial in a near ideal situation) T is accepted there is need for caution. The tendency for educators to rush the move from conception (conjecture) of an idea to implementation is widespread. The proposal to introduce an ideal situation trial between the two will achieve little unless at least two concomitant conditions prevail. First T^1 cannot in educational research be a crucial test. Extensive replica-

tion is essential and the fact that identical replications are not possible is an advantage, not a disadvantage. Only to the degree that T can be subjected to a range of different, but equally ideal trial situations as to design etc., will the weaknesses be revealed and the H refined and modified. There has been neither test nor trial unless H has been modified in some way. Second, T² which is the beginning of practical utilization cannot be condoned until substantial progress at the T¹ stage through a series of replications has been made. We need to think in terms of a series T¹, T₂¹, T_n¹ and advance to T² only when n is deemed to be great enough to justify generalization of T². This necessitates restraining the enthusiasm of innovators and that is not likely to be readily accepted by most practitioner theorists.

Perhaps the most important (certainly the most neglected phase) of T and of the model as a whole is the non-ideal situation phase which, for simplicity, is designated as T². Any technique, method, appliance, or machine has value proportionate to its validity or usefulness under normal operating conditions. The airliner that can only be operated effectively and economically by specialized personnel not available under *normal service conditions* is of little value. The educational innovation that can only be made to work effectively with *highly specialized personnel under ideal conditions* is of little value. Again we should note that the history of educational method from the critical questioning techniques of Socrates to the freedom of A.S. Neill is a source of examples of innovations introduced because they were highly successful in the contexts created by their developers. The same history reveals that in the hands of others they proved much less valid. About this at least two things must be said.

First there is a sense in which it probably is the case that there is a subjective personality related dimension to the teaching and the educating act. It follows that if this dimension is such that its retention is essential to the success of the act then that act cannot be performed by others unless that dimension can be developed or translated. Therefore one of the main tasks of both T¹ and T² is to determine (a) if such dimensions exists, and (b) if they can be developed by or in others. If (a) is established and (b) is not achieved then the innovation must be rejected as being inappropriate for general use. What might of course be done would be to identify the traits necessary and the personnel who exemplify them and confine its use to such personnel. One might note parenthetically that techniques such as team teaching are not effective with all teachers.

Second, there is a large number (probably the majority) of cases where there are no necessary unique subjective personality traits, but it does not follow that no modification of post T¹ versions of H₁ are required. The determination of the form and extent of these modifications is partly the concern of T¹ but more the responsibility of T².

The research phase T² is conceived as a form of action research—research carried on by practitioners in their day to day practice—and at least some educators today are advocating encouragement of this type. For reasons which, directly or indirectly, have been referred to in the foregoing, there is a case to be made. Certainly a very strong case can be made for research in the classroom by those responsible for the classroom as a means of disseminating research findings to the classroom. There are com-

plex problems as well as other advantages that are worthy of consideration. The problems fall into two general categories: (i) the lack of adequate facilities and skills by classroom personnel, and (ii) the related problems of conceptual orientation of these personnel. The solution to (i) involves considerable training and selection. The solution to (ii) is even more complex. In any practice—medicine, law, engineering, teaching, etc.—it is not realistic to expect the practitioner to be conceptually oriented toward research to the degree necessary for sustained independent operation.

The aspect of educational research that is of concern here is that which has a direct bearing on educational development in the context of teaching and learning styles and strategies. It differs significantly from any form of research the end product of which is cognitive in the “knowing that” sense. Basic social and educational research are concerned with getting to know the characteristics of human learning and human development. Our concern is getting to know how to most effectively incorporate and use basic knowledge but it does not follow that basic knowledge will not accrue.

The remarks in the foregoing description of the model were identified with arrow (a) in the diagram. They have therefore followed sequentially through S, H and T^1/T^2 in a logical order which is too much akin to a simplistic interpretation of that nebulous thing called scientific method. Arrows (b) and (c) are incorporated to indicate awareness that this order may not be valid or viable. One can readily conceive a host of situations wherein hypothetical H_1 or T_1 moves may be necessary before any progress with respect to systematic organization of the allowable elements of a conjecture can be made.

The dual direction interconnecting arrows among S, H and T are designed to indicate the intrinsic relatedness of these aspects. One might think of the whole unit as tri-cellular with each cell developing in part independently but not without relatedness to the others. The biological fertilization analogy may not be inappropriate particularly if one considers the separate cases of S, H and T being either separately energized or the results of a single energization and subsequent splitting.

Conjecture and Refinement

The Popperian principle of conjecture and progressive refinement of the initial idea is inherent not only in the C phase but in the whole model. A slightly more detailed consideration of that principle is necessary but first of all one point must be made. Karl Popper, operating in the context of critical epistemology specific to scientific discovery and set upon the task of eliminating subjectivism, can meaningfully speak of refutation. His is a context in which hypotheses, in principle at least, can be formulated with sufficient precision (riskiness) and measurements and experiments carried out with sufficient accuracy and control respectively to make refutation a feasible thesis. But even so, such refutations are seldom achieved and it is noteworthy that Popper, in later formulations of the principle (1972a, p. 121), refers to error elimination rather than to refutation.

Popper's clearest formulation of the principle is expressed in the formula

$$P_1 \rightarrow TS \rightarrow EE \rightarrow P_2$$

where P_1 is the initial problem, TS, tentative solutions, EE, resultant error

there exists a real function and a real place. The second arises out of the first. The positive assignment of a function necessarily implies availability of personnel qualified to perform. At this time due partly to the inadequacy of practitioner preparation programs and partly to their subsequent non-involvement in, and isolation from research, few practitioners are qualified. Those that have shown interest and become competent have "graduated" from the practitioner role. From this it follows that the organization of research and the design of its venue must provide means for extension of practitioner competency in a way that does not lead to the cessation of their performance as practitioners.

This can be achieved by applying the cyclical concept of the theoretical model to the staffing pattern of research centres. Research centres can be conceived as the nucleus as to both function and personnel. This means that they will house permanently only the facilities that are specialized and require a permanent location (computers, libraries, etc.) and those personnel who are specialized (programmers, statisticians, theorists, technicians) and who contribute to the centre's activities in a wide range of projects.

The key to facilitating the capacity of research to influence practice significantly is the T^2 stage. Its integration into the research program can be achieved by organizing continuity of research activity and teaching practice. This can be done, the writer submits, by seconding selected teachers to work as researchers in research centres during the T^1 phase of a project. Upon completion of an appropriate period of secondment determined by the progress of the project or by the competency of the seconded teacher-researcher, they would progress to functioning in the project at the T^2 stage. This procedure opens the way not only to the development of test by trial in ordinary classroom situations but also to the development through in-service training of competent teacher-researchers.

The effectiveness of short term secondment of teachers to engage in teacher education program on a basis operationally equivalent to permanent faculty has been well established at Simon Fraser University. The administrative pattern developed is readily applicable to research activities and the benefits to be derived by teachers and by research may be very great. At the very least the scheme will locate in schools and classrooms a nucleus of research oriented teachers. As such it might be viewed as a valuable form of inservice education.

One final note must be made. Research activity is not facilitated by static organization or conceptualization. It may well be the case that the most productive research unit is one in which not only practitioner but also the more narrowly technical personnel are seconded for short terms and caused to flow into and out of other areas of activity.

Educational Research: Concepts and Models

References

- Broudy, H. S., Ennis, R. H., & Krimerman, L. I. *Philosophy of educational research*. New York: John Wiley & Sons, Inc., 1973.
- Educational Theory*, 1973, 23 (2), pp. 159-176.
- Kerlinger, F. N. The mythology of educational research: The methods approach. In H. S. Broudy, R. H. Ennis & L. I. Krimerman (Eds.) *Philosophy of educational research*. New York: John Wiley & Sons Inc., 1973.
- Philosophical Redirection of Education Research*. Seventy-first year book of the National Society for the Study of Education, Part 1, 1972.
- Popper, K. R. Epistemology without a knowing subject. In *Objective knowledge*. London: Oxford University Press, 1972.(a)
- Popper, K. R. Of clouds and clocks. *Objective knowledge*. London: Oxford University Press, 1972(b).

D. FRIESEN

The University of Alberta

A Study of a Public School System Hessen, West Germany

This paper summarizes some of my findings on the organization and administration of public education in Hessen, West Germany. The data, collected during a four-month period, resulted from visits to schools, interviews and examination of documents of various kinds.

The positive characteristics of public education in Hessen are readily apparent to the visitor. Among others they are the strong academic basis, the early writing and reading programs, the emphasis on excellence, the pragmatic vocational programs, and the development of responsibility among students. The focus of this investigation is only peripherally on these characteristics; it is mainly on the organization and administration of public schools, and on proposed changes in these areas.

To achieve these objectives the paper reviews briefly the historical background of public education in Germany, then discusses the current educational structures in Hessen. This is followed by a more detailed analysis of the administrative structures in the West German state. The final section of the paper deals with educational reform in Germany, both in terms of objectives of the reform and in difficulties encountered in the realization of desired changes. (Dr. Friesen is Professor in the Department of Educational Administration, The University of Alberta.)

Historical Perspective

A brief reference to the historical development of education in West Germany is essential in order to comprehend more fully the complexity of current educational problems in that country. There is a proud tradition, as in the case of most European nations, lodged in the Gymnasium, the humanistic and scientific curriculum, the exclusiveness of certain schools and universities, and the customs and practices related to each of them.

Control of Schools

In early days, schools were predominantly church controlled; priests and monks were the undisputed decisions makers in educational matters. According to Butts (1955, pp. 199-201; 359-364) a long period of controversy,

dating back to at least 1598, raged, during which time the state gradually assumed more and more control over schools. General formulas were thus developed which resulted in more specific laws such as the one in Prussia (1794) which stated that schools and universities are institutions of the state and that all public schools and educational institutions are under state control. But throughout the nineteenth century the state control of schools did not fully materialize. Even after the externa had been absorbed by the state, it took decades before the interna were removed from church control.

It is no wonder that, when after the first World War, the complete break was to be made toward a state controlled democratic school system the change could not be effected immediately. Many influences lingered on and these became powerful forces when change was perceived to threaten the traditional educational system. Many of the organization arrangements and school practices in German education can be traced back to the revolutionary change in education accomplished in the Humboldt era. In many quarters, German educators, politicians, and students are questioning the appropriateness of an educational system from a previous century in a time when technology has affected almost every aspect of life.

Control of Curriculum

Traditionally, curriculum was a matter for the priests, monks and scholars. The classics, religion, languages and mathematics were the undisputed disciplines for the students in the Gymnasium. Later the sciences were grudgingly admitted to occupy a small part of the curriculum.

In spite of changes in technology, social order and ideologies, the curriculum of the Gymnasium follows its traditional course. The state may control the decision making processes, but it does not or will not interfere with the interna of the schools. De facto curriculum decisions are still made outside the arena of student, parent, or state activity. Just as in England, the teacher in Germany is free to make many curriculum decisions. But in actual fact he makes few that are not those handed down to him. Conformity lies not in the law but in the historical precedent or tradition.

Pressures on the Schools

Following the period of idealism during the early nineteenth century, Germany experienced a reaction that slowed the progress of school development. However, the basic structures of the schools were established. The pressures of an industrial nation became more noticeable. Vocational schools received much attention and these gradually became important educational institutions in Germany.

The German school system, perhaps because it never completely severed its church connection, continued to feel the influence of ideologies. This influence appears to have increased dramatically of late. Together with this development have emerged hopeful signs of a reawakening of the importance of education in a democratic nation.

Critics of German education are voicing concerns that control of schools is ambiguous and probably results in inefficiency and uncoordinated planning, that administration of schools is bureaucratic and unplanned, that curriculum of schools is accepted uncritically, that in content of courses, in-

dustry and society are greater determiners of education than the individual, that education, emerging as one of the most important concerns of the state, is refusing to change its century old orientation to organization, control and development, and that current structures in education retain the elitist tradition and as a result disregard demands for greater equality of opportunity.

An international study (Husen, 1967, p. 114) of twelve countries sought to document the relationship between school structure and the degree of equality of opportunity in education. An index of the proportion of middle class children in schools was developed; the larger the index the greater would be the proportion of students from the middle classes during the years of schooling. The larger the index the more unfavorable was the school organization to the working class children. The index showed that for mathematics students the U.S.A. (1.9), Sweden (2.1), Israel, Belgium (3.6), Australia (4.7), Japan (6.0), and Finland (6.0) had the lowest values. The three with the greatest values were England (16.2), France (17.3), and West Germany (45.3).

Clearly, according to these findings, countries with comprehensive schools have attained much greater equality of educational opportunity than those with a tiered system, where streaming of students is practised. The study also showed that the comprehensive schools were more efficient in that they led more students to successful completion of high school, without any loss in performance of the top students. Similar results have been found in recent IEA studies (Tanner & Tanner, 1974).

Current Educational Structures in Hessen

Schools

The following descriptive summary of the Hessen schools is based on personal visits, inspection, interviews and publications by The Bildungskommission des Deutschen Bildungsrates (West Germany's Education Commission) (1970, 1973), Hauke (1970), Roehrs (1969), Rolff (1970), Hentig (1970), Brock (1972) and others.

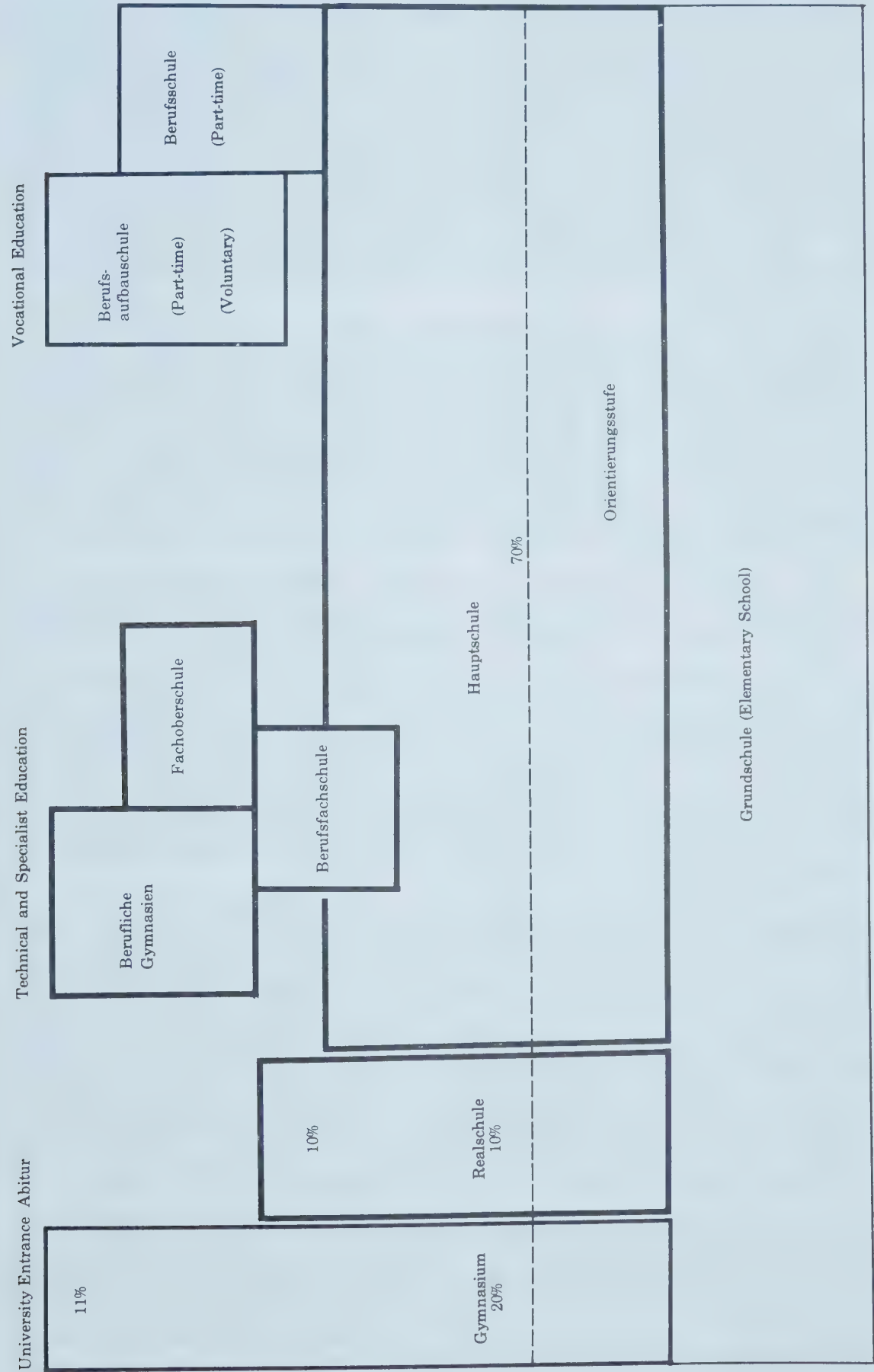
(a) *Vertical Structures*. (Figure 1) The basic elementary school in Hessen, as in other German states, is the *Grundschule*.

Since 1918 all children regardless of background or social class must attend this school for four years beginning at age 6. This school was designed to break the educational monopoly of the rich or elite, and lead to a democratization of education in Germany, and to greater equality of educational opportunity.

Brock claims (1972, p. 20) that the aims were not achieved for all children and that the inequality was merely veiled. The National Socialists did not change much in education except to remove the confessional nature of the schools. After the second World War no reform set in, except the restoration of the confessional or religious program and the deletion of the Nazi ideology. The manifest goal of democratization still plays a secondary role to the latent goal of selection.

Following the four years of comprehensive schooling the majority of children are streamed into the *Hauptschule* a five year secondary school. The students are mostly those of the working class, who are at this early

FIGURE 1. THE SCHOOL STRUCTURES TODAY—HESSEN, WEST GERMANY



stage clearly differentiated from the other classes. Many writers and educators claim these are the weakest schools in the system attributing this weakness to curriculum, teacher qualifications, high student-teacher ratio, and the low esteem held for the Hauptschule leaving certificate.

About ten percent of all students are streamed into the *Realschule*, a school designed to prepare students for middle to higher occupational levels. The curriculum is a mixture of practical and theoretical activities defended by industry and by Realschule teachers who are afraid of losing their somewhat higher status than that of Hauptschule teachers (Brock, p. 27). The school stands between the Hauptschule and the Gymnasium, and thus obtains support for its existence since it helps to enhance the status of the Gymnasium.

About 20 percent of all students enter the *Gymnasium* (Deutscher Bildungsrat, 1970), the elite school in the system, and about 11 percent complete the *Abitur* (Brock, 1972, p. 29). Boys are overrepresented; the working class is seriously underrepresented. Entrance is based on the recommendation of the Grundschule. The curriculum is strongly oriented toward languages, mathematics, and science. The *Abitur* leads to university or professional training. Evening schools both for the Realschule and Gymnasium are provided for students who wish to move toward the Abitur on a part time basis.

The schools of Germany are clearly differentiated, following the Grundschule, into three distinct and separate streams. The single educational ladder ends after the fourth year of school.

(b) *Early Selection*. One very obvious barrier toward democratization of the school system is found in the early streaming of children. Some parents, even before their children attend school, have their children drilled to "Make the test." During the Grundschule the democratization aims are subverted with the overarching fact that a decision is pending which will determine to a large degree the life-style of the child. The German mind seems to be able to live comfortably with this educationally questionable practice.

(c) *Curriculum*. Most educational books in Germany that discuss curriculum tend to focus primarily on that of the Gymnasium. This curriculum is viewed generally as untouchable. Every child must study languages, especially Latin and one or more of French and English. In the Hauptschule the curriculum is more practical and job oriented, while that of the Realschule is both practice and theory oriented. A very recent development in the Hessen educational scene is the re-writing of curricular guidelines in terms of educational objectives. The *Rahmenrichtlinien* (curriculum guidelines) have generated considerable interest and debate in various arenas particularly the political. As an example, a six hour marathon panel discussion was shown on television on the subject.

(d) *The School Day*. The half day school from 8:00 a.m. to 11:30 a.m. for the Grundschule and from 8:00 a.m. to 12:30 p.m. or 1:00 p.m. for upper years is still defended and practised widely. This means that children always have homework, and recitation occurs in school. There are some suggestions to change to a five day week with a full school day which might permit new modes of instruction and which would be in harmony

with the five day work week now becoming general in Germany. But the change seems too threatening to teachers and Schulraete (administrators) and even to parents.

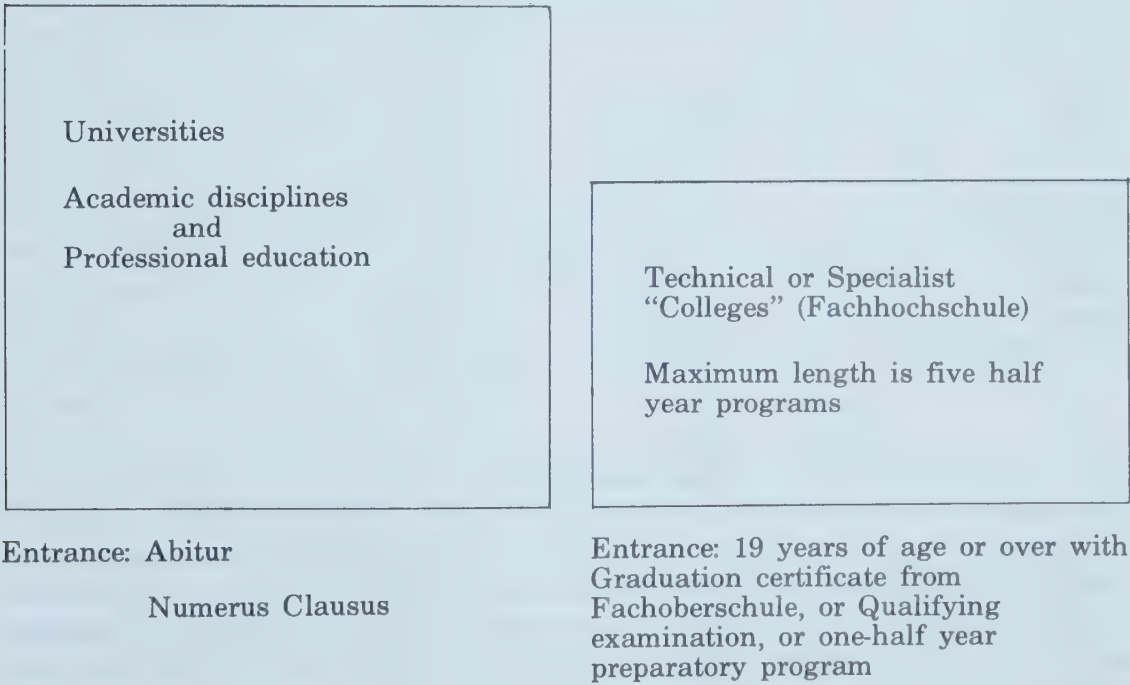
(e) *Abitur*. From early schooling German education is Abitur dominated. Abitur refers to the successful completion of examinations following Gymnasium. It is interesting to note that most German critics of education refrain from attacking the institution of Abitur. Hentig (1970, p. 179) comes close to it when he advocates several changes in the Abitur. The institution is deeply imbedded in German thinking—"Doch!" "Ja, Ja, so ist es doch." "Natuerlich", "Eben." For an outsider the Abitur represents symbolically the essence of the elitist tradition in German education.

Higher Education

Figure 2 summarizes briefly the educational opportunities open to German students beyond the Gymnasium or beyond the Fachoberschule or Berufsgymnasium (See Figure 1). Recently Numerus Clausus has been added to the Abitur in the selection of university students. In 1973 only 15,864 West German students out of 52,177 were awarded a place for the course of study of their choice (The Times, 1973, p. 9).

FIGURE 2. HIGHER EDUCATION—HESSEN, WEST GERMANY

State examinations for teachers
conducted by Department of
Education following 4 to 6 years study

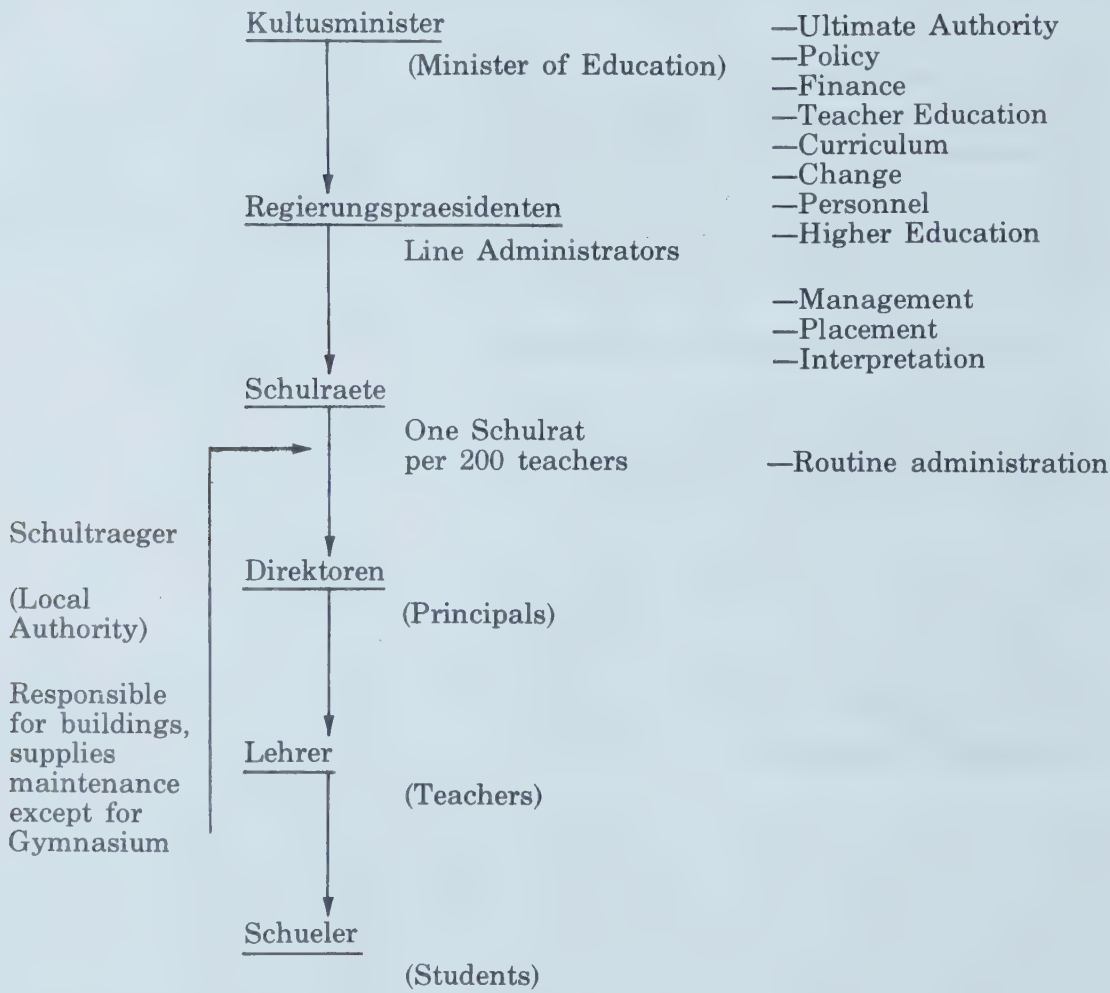


Entrance to the Fachhochschulen is possible for students at least 19 years of age who possess a Hochschulreife certificate (See Figure 1). They can also enter by passing an entrance examination or complete a half year preparatory course.

Administrative Structures in Education

With the long history of gradual development of the present school system, complicated and varied administrative structures have developed concomitantly. The following descriptions, inadequate as they may be, may help to provide an understanding of the processes of governance in German education. The sources again are from the references mentioned and the interviews held with various officials and researchers.

FIGURE 3. ORGANIZATION FOR SCHOOL ADMINISTRATION IN HESSEN, WEST GERMANY



Figures 3 and 4 provide an oversimplified picture of the administrative structures. In Parsons terms the Kultusministerium (Minister of Education) would represent the institutional or political level. The next three positions would be the management or administrative level, while the last three positions would be the operative level.

A number of additional institutions or organizations are influential and/or operative in education affecting the work that is done at the operative level, the routine of the management level, and the policy at the political level. Five of these must be elaborated upon briefly to make the presentation of school administration more representative of the actual processes as they occur.

FIGURE 4. SUMMARY OF INFLUENCE ON SCHOOL ORGANIZATION AND ADMINISTRATION IN HESSEN

West Germany—

no constitutional rights in education yet “de facto” influence is felt through recommendations, participation, financing of higher education, and control of vocational education.

In addition, four super regional educational bodies exist. They are limited to recommendations.

1. Standing Committee of Ministers of Education for planning, coordination and exchange of ideas.
2. Commission on Educational Planning.
3. The West German Education Commission—conceptualize.
4. “Bundesministerium” for Educational Science—research.

Parent Councils

- Input possible at all levels
- No authority for decisions

Teachers Councils

- Have authority for decisions at school level
- Not used frequently

Student Councils

- Input possible at school level
- No authority for decisions

The Federal Republic. The federal government of West Germany has no constitutional rights in education; the rights are vested in the state (Land). Yet “defacto” influence occurs in several ways. First, it derives from the overall policies pursued by the West German government and by the rights guaranteed by it. Among these latter are the guarantees that everyone has the right for his free personal development (Bildungsrat, 1970, p. 25), and the right of a free choice of occupation. Second, it exerts influence by recommending or supporting recommendations in education. Third, it actually participates in planning commissions, and finally it contributes financially to higher education and to research institutes.

Specifically, there are four super regional educational bodies that influence state decisions in education.

The Standing Committee of Ministers of Education is involved with planning, coordination and exchange of ideas. The Commission on Educational Planning is composed of representatives from the states, and also from the Federal Republic. The German Educational Council is another body engaged in conceptualizing educational plans. Finally, there is the Federal Ministry for Educational Science which is engaged in research and development. All of these bodies can only recommend action or policy to the states, they cannot go beyond. Their recommendations are widely disseminated and are influential.

Schultraeger. (Local authority for buildings and supplies). The provision and maintenance of school facilities, equipment and supplies is vested in a local authority. These authorities can be cities, divisions or county-like units. The Schultraeger vary for different parts of the school system. In general the Schultraeger for all public schools are local authorities, but for

the Gymnasium it is the Kultusministerium (state department of education).

A school director is thus subject to a dual authority. For staff and policy problems he contacts the *Schulraete* while for material problems he contacts the *Schultraeger*. In some places like Frankfurt am Main the *Schulrat* is responsible to the Kultusministerium for the *interna* (staffing, program) and to the local authority for the *externa* (buildings, supplies).

Parents. Parents have their own formal organization which can act as a lobby at any level of the hierarchy. The parents can thus discuss problems of curriculum, teachers, schools, governance or policy. The decision-making is outside their prescribed rights, theirs is influence alone. The effectiveness of the influence is clearly related to the person not the position. Recently some of the "Landeselternbeirate" have exerted considerable influence, and the minister is using this structure to involve parents in discussion and decision making on curricular matters.

Teachers. Teachers have by law the right to determine a number of matters through the *Konferenzrecht* (rights granted by law). However, this is not used too much and teachers, as a result, may be influential mostly at the operative level. The fact that teachers are civil servants, and their progression up the hierarchical scale is determined in part by their superiors may deter them from aggressive participation in decision making.

Teachers are also active through their own associations which generally tend to support innovative thrusts in education. The actual effort is carried out through publications, workshops and cooperation with the minister in innovative programs.

Several obvious difficulties make the work of teachers' associations less effective than could be expected. First, only about one-third of Germany's teachers are members of the *Gewerkschaft* (teachers' association). Even though the membership is increasing, it will take time before the association (union) can speak effectively for teachers in Germany, and until such time the major thrust of the association may have to be with teachers rather than with policy makers. However, the leaders of the teacher associations are moving into the politics of education, and since they don't appear to be in conflict with the minister on many issues, may exert some influence.

Another difficulty is related to the historical development of teacher education in Germany. Teachers of the Gymnasium have received their education in universities and have continued to maintain their own association (*Philologenverband*). The members of this association are more inclined to resist changes in educational structures and curricula. They tend to have close ties with universities and university graduates who find their way into politics, law, or other influential positions. Many of these also move into middle management positions or higher education and thus are involved in the political arena and decision making processes. Teachers of the *Hauptschule* and usually of the *Grundschule* were until recently not university educated. This cleavage, even though all are university trained now, still creates communication problems and, probably more serious problems of conflicting goals, values, and priorities.

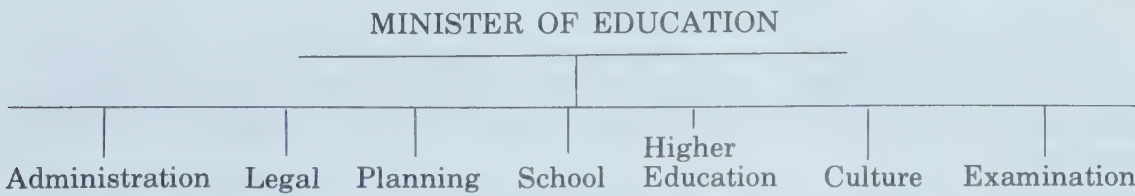
The third concern about teacher influence relates to the way that

change and innovation have become linked to the radical or leftist teacher. This tendency has resulted in a polarization of opinion where one group wants to change slowly, carefully and according to goals and plans, while the other wants to move rapidly to replace what in its words is an “im-moral” system. This polarization has helped to slow down meaningful change. There are a few educators and politicians who are making themselves credible by avoiding the association of extremists and building their cases for change on rational, empirical and humanistic grounds, almost free of ideology or reaction. Though these are in the minority, they represent a promising thrust in German education.

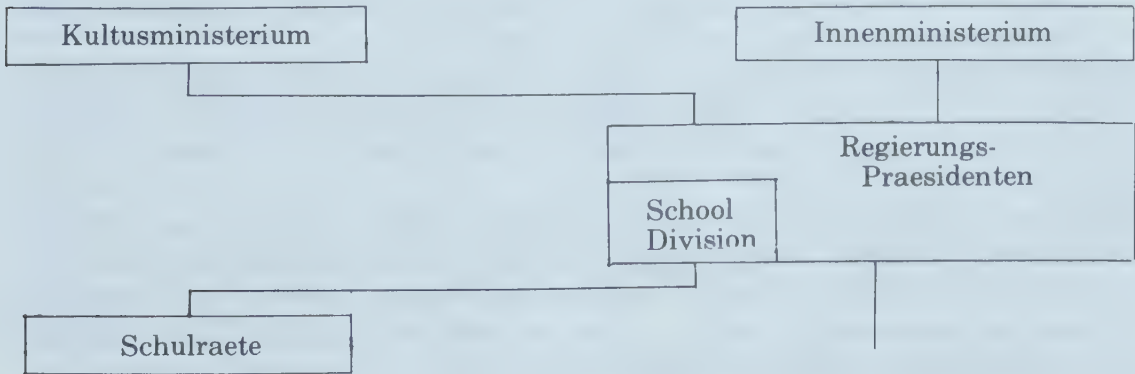
Universities and Research Institutes. By their own admission, members of the educational and research institutes have minimal direct influence on educational reform. However, both theory and research are actively in-jected into the educational discussions at all levels by these institutions and thus they may make significant contributions. In addition many university professors, researchers, and writers are actively involved in educational reform proposals. There are indications though that educational reform will result more from political moves than from educational theory and research. It is also interesting to note the com-promises researchers seem prepared to make especially in comparative research. They are always ready to state that Germany must avoid the errors made by other countries, that is why they pursue research in other systems.

The basic line sketch of the administrative structures for education in Hessen, West Germany as seen in figures 3 and 4 shows the numerous in-fluences on the decision making processes that occur.

The Kultusministerium has the ultimate authority in education. The minister and his staff are responsible for policy, for finance, for change and innovation, for curriculum and examinations, and for legal and ad-ministrative matters. At the highest level the Kultusministerium is struc-tured something like the following:



A confusing structural aspect occurs in connection with the Regierungs Praesidenten as shown below:



Research has shown that jurists are represented fairly strongly at the higher levels of the administrative hierarchy (Muller, Vogelsand, & Bessoth, 1973, p. 76). Some 40 percent of the top educational officials are from the legal profession.

The educators at the top level arrive only late in life and thus may come with a bureaucratic career orientation rather than a creative desire for change. The lawyers, on the other hand, come equipped with legal ideas and are probably less interested in educational reform than in maintaining an efficient and traditional system of education.

This brief sketch illustrates the complexity of the administrative structures in German education. The analysis can easily lead to questions on efficiency, planning, accountability, supervision, and motivation in the system.

The rationalization of the administrative structures seems imperative. One way would be to study the functions at each of the three administrative levels and to examine ways in which the structures could be reordered for the improvement of the system. During interviews with various incumbents in the administrative structures much emphasis was found on day to day operations and on global issues such as the necessity of radical reform or the danger of radical reform. There was sharp criticism on both the advocates and the resisters of radical change. All seemed to agree on global goals, but little agreement seemed apparent on means of arriving at these goals. Thus the upper, the political and policy making arena was active in proposals and debates, while the lowest arena was active in day to day activity. However, middle management seemed absorbed in translating orders from above to those below. Little attention was apparent on such functions as in-service education, curriculum development, supervision, motivation of teachers, evaluation beyond the formal practices, or leadership in change and improvement. The functions of middle management seemed lost in the daily routine of staffing, examining and carrying out directives.

Educational Reform in Germany

Before examining current changes that are proposed and to a degree implemented, it is helpful to review salient characteristics of the educational system. These are the characteristics around which the change controversy revolves.

A Summary of Characteristics of German Public Education

1. Though the state is the sole authority in education, the Church has retained an influence through confessional schools (curriculum) and through taxes going to state churches.
2. Early selection of the more able students continues to be generally practised in Germany. Even though this selection may be delayed somewhat beyond the fourth school year, it is still not questioned too much.
3. The more able students are required to follow a curriculum which is narrowly defined and rigorous. The curriculum of the Gymnasium is a humanistic and scientific curriculum, that of the Realschule is practical-theoretical, while that of the Hauptschule is "einfach, volkstuemlich und genuethaft" (Brock, 35).

4. Student participation in decision making is relatively rare.
5. The *Abitur*, the examination following Gymnasium, remains a dominant factor in the student's educational plans and curriculum.
6. Bureaucratic arrangements are hardly questioned in regards to the Gymnasium and Realschule.
7. Half-day, six day a week, schools are still strongly supported by parents, teachers and politicians.
8. Different financial arrangements exist for the different school types.
9. The clear-cut three ladder system of education is accepted and generally adhered to (Brock, 1972, p. 34).
10. Administration of schools is interpreted as following traditional and legal directives, resulting in a tendency toward the authoritarian personality in administration.

Change in German Education

Pressures for change have been increasing in German education. The changes have been directed toward seven main aims:

- (1) Greater equality in education (Deutscher Bildungsrat, 1970, pp. 12, 30, 98-213; Hentig, 1970, p. 22; Rolff, 1970, p. 22; Deutscher Bildungsrat, 1973, p. 2)
- (2) Greater attention to newer teaching-learning strategies (Deutscher Bildungsrat, 1970, pp. 36-88, 40-95; Hentig, 1970; Hauke, 1970, p. 22; Deutscher Bildungsrat, 1973)
- (3) Greater sensitivity to needs in society (Deutscher Bildungsrat, 1970, p. 12; Hentig, 1970, p. 22)
- (4) Greater participation in decision making (Geizler, 1970, pp. 53-145; Deutscher Bildungsrat, 1973, pp. 4-14)
- (5) More attention to system planning (Rolff, 1970, p. 56; Jensen, 1970; Edding, 1970).
- (6) New teacher training programs (Deutscher Bildungsrat, 1970, pp. 215-252)
- (7) Special education for administrators (Deutscher Bildungsrat, 1973).

The response to each of these reform objectives has begun to be formulated in specific change. Some illustrations follow:

Equality in Educational Opportunity

The demand for greater equality in education led to a re-examination of school organization. The result was the proposal of a *Gesamtschule* (Comprehensive school). Figure 5 presents the model of the school which has been adopted in a number of districts. Hessen claims about 75 *Gesamtschulen*, some of which are fairly well integrated; others have simply brought the three streams of students under one roof. The proposals stress a core curriculum of social studies, science, German, mathematics, a foreign language, sport, religion and art in the second stage of school (years 5-10). In addition the student would have to select at least one course from a specified list and be able to choose from others for additional courses. For the last 2 or 3 years of school the core would be reduced and the range for selection increased.

For many the *Gesamtschule* is a politically inspired threat to the old

FIGURE 5. GESAMTSCHULE (COMPREHENSIVE SCHOOL)

SEKUNDARSTUFE 2 (Secondary Phase 2) (2-3 years)
Choice of basic and practical courses

Language	Sport	An offering outside of all subject areas, theory and practice
S. S. (economics, politics)	Music	
Science	Religion	
Mathematics	Religion	

SEKUNDARSTUFE 1 (Secondary Phase 1) (5-7 years)

<i>Pflichtunterricht (compulsory)</i>	<i>Choice</i>	<i>Additional</i>
S. S.	2 or 3 foreign languages	(One must be selected)
Science	Technical courses	Practical
Mathematics	Business	Speech
Foreign Language	Social	Therapeutic
Sport		Training
Religion		
Art		

PRIMARSTUFE (3-5 years)
(Primary Phase)

Kindergarten - Vorschule - Duration 2-3 years

system; for others it is a partial answer to the need for greater equality of educational opportunity.

Teaching-Learning Strategies

The demand for the introduction of newer teaching-learning strategies has been approached at two levels. The teaching strategies advocated are alternatives to the German teaching recitation school period with the required homework during the afternoon. To break this tradition full-day school is being implemented in some places. This will permit greater attention to such things as school centered learning activities, group discussion, audio-visual and other aids, and team teaching. This move is also supported through the new school building designs which permit much greater flexibility in organizing the teaching-learning activities. However, the change is coming only very slowly.

Some teachers openly resist the full day school because it would change the teaching-learning practices of German education. Homework and recitation are seen by these teachers as essential in the learning process. There is some humor involved in the discussion of this topic in German circles. Of course no one would want to take the afternoon nap from the teachers. Teachers are quick to point to the long work week they have in the present system and the dangers of the extra burdens for teachers and pupils if the full day school were introduced.

Interesting too is the parent opposition to the full school day. So deeply ingrained is the tradition of the half-day school that parents can't see how it would be possible to have less homework time for their children. Administrators and politicians point to the need for more space, for

lunchroom, for more personnel, more supervision and the like if a full day school were implemented. The cost would be prohibitive. Against all these arguments are those of some educators who would like to try new forms of organizing for learning and those parents on a five day week who would like to have the week-end free and to have their children in school during their work days.

The German elementary school week consists of about 18-24 hours compared to about 30 hours in Canada and about 35 hours in the British primary school.

The second response in the area of newer teaching-learning strategies relates to curriculum change. An example of this is found in the hotly debated *Rahmenrichtlinien* of Hessen (Curriculum guidelines of Hessen). These guidelines have been developed to fill a vacuum created by the attention to organizational reform and were prompted by new aims in education and new educational practices in achieving these aims. Accordingly the guidelines were based on establishing learning objectives in considerable detail and providing latitude for the teacher in planning activities and experiences to move learners towards these objectives.

The fact that these guidelines moved away from subject matter or content goals to process goals led to much criticism and discussion. It also led to the consequence that school centered curriculum development must accompany the change. No detailed content was prescribed any more. This brought with it the need for different kinds of teacher education, of in-service education, and of a new type of administrator who could be a leader, a stimulator, and a colleague.

The general aims of the new guidelines are to equip the student to participate actively in a democratic society by developing skills in communication, decision making, and evaluating, and criticizing analysis of social institutions. So threatening is this change that it is attacked by prominent politicians as "Education for class struggle." In the hands of some educators this may become reality, but the general support is based on the argument that the pluralism in a democratic society calls for citizens trained in conflict resolution, citizens who are capable of evaluating, communicating, participating and decision making on their own. A feeling exists that educational reforms of the present administration in Hessen will be a major issue in next years' state elections.

Attention to Needs

Both the new organization and the new curriculum of the schools are to reflect the school's sensitivity to societal needs stemming from the advanced technological world and the emerging social patterns.

Hentig (1970, p. 22) claims that the Gesamtschule is a concept that one must support or oppose in principle and let it work out the technicalities, because it permits the school to adapt to such things as technological needs, the social needs, and the rapidly changing social structures.

The sensitivity to societal needs is clearly revealed in most of the documentation supporting the Gesamtschule and the new curriculum guidelines. The sensitivity to the needs of the individual student has not been as clearly formulated even though the motivating factors are equality of educational opportunity, development of the individual, and freedom to

choose an occupation. But the approach reflects more clearly the changing needs of society than that of the individual. In all probability the second phase of this German attempt to democratize its school system will be motivated more by concerns of the individual student and will call on participation, greater choice, and humanization of education.

Participation in Decision Making

Greater participation in decision making in schools has received very little attention. The Bildungsrat (1973) raises this as a serious issue.

Though direct participation has been limited there is a political participation which goes something like this. The Kultusministerium will develop a plan which when accepted moves down the hierarchy with little question. If resistance is found it is sent to the bottom of the hierarchy where students, parents, and groups in the community enter into confrontation types of discussions. In the meantime the administrators and teachers appear more as reactors than leaders in the discussion of such ideas.

The latest set of recommendations from the Deutscher Bildungsrat, though concerned primarily with the reform of organization and administration of schools, comes out strongly for independence of schools within Departmental guidelines. This would greatly increase teacher, student, and parent participation in decision making (Deutscher Bildungsrat, Part 1, 1973). Participation will be discussed in greater detail in the section dealing with administrative reform.

System Planning

More attention is gradually given to system or nationwide planning. Rolff (1970) advocates a continuous (rollende) reform through educational planning. Preissler (1968) illustrates a model of planning in the region of Kassel.

Just as in most countries of the world, education has become one of the critical issues in West German society. The crisis as documented by OECD and UNESCO has the following distinguishable characteristics:

(1) The explosion in numbers of students seeking further education is present. Shortages in space and teachers result. Children from all social classes must be admitted.

(2) The shortage of resources seems to increase continually.

(3) An increased demand for higher education results in the increase in cost per student and the spread of shortages to all levels of education.

(4) There is a resistance to modernize educational methodology which leads to a discrepancy of the school as a sub-system of a rapidly changing system.

(5) Education does not meet the needs of society nor of individuals in a satisfactory manner. It does not provide the foundations for the free development of the individual in a rapidly changing world.

Theoretically, the designs for planning have been given considerable attention. The actual planning programs of the states (Laender) are much more simply conceptualized than the ideal type of planning. The following points illustrate a planning pattern which has been proposed:

1. The current situation
2. The goals of education—individual, societal

3. The resources or materials needed

4. Efficiency.

There seems to be only a very modest attention to systematic planning, evaluation, to the integration of information from the local setting and the scientific world, and to the more scientific problem solving procedures.

The tendency exists to lapse into a political, only partly rational, approach to planning new educational structures in Germany, as in most other countries. Perhaps this leads to change in peripheral matters rather than in more central issues.

Teacher Training

New teacher training programs are advocated, but as yet changes have been few. Since teachers for the different school types receive different kinds of education, the problem of revising teacher education is complex and difficult. The difficulty is increased because of the current shortages of teachers and perhaps also because the present system contains a status hierarchy that may be threatened with change.

However, modern teaching methods are being adopted by some teachers, especially in the Gesamtschule. With further organizational change toward a full school day, teachers will probably, through in-service and university training, adapt to the new needs. Newer teaching technologies are somewhat alien in the old system.

As with so many proposals for change, the proposal for a change in teacher education faces great difficulties because no move is made to remove the major impediments for such a change. In order to facilitate the acceptance of new teacher education programs, some changes need to be considered in several traditional practices.

The present training program runs from 4 to 6 years and is culminated by a final examination. This examination, though carried out by university professors, is controlled by the Kultusministerium. The teacher then must teach and study for two years after which he faces his second examination, this time carried out and controlled by the Examination Board (Prüfungsamt) of the Kultusministerium. The tendency is to move toward traditional expectations to satisfy the examiners rather than toward innovative ideas that come from sources outside the system.

The second impediment toward change in teacher education lies in the teacher hierarchy and promotion system. After the initial two years of provisional status, the candidate, upon successful completion of his second examination, is known as a subject teacher and moves through the status-payment categories A-9 and A-10. Then he is known as a teacher who can teach in the Realschule and moves through status-payment categories A-11 and A-12. A-13 is the next category for the Studienrat which is followed by A-14, the Oberstudienrat. The latter two categories refer to teachers in the Gymnasium and to teachers who have some administrative responsibility.

These teachers categories are followed by A-13 to A-16 for teachers moving into administrative positions (Schulräte) with similar status-payment classifications. These are followed by H-1 to H-4 categories for those in higher positions. All of the above statements are approximate, but they illustrate the existing hierarchy.

Added to this is the fact that all teachers are civil servants and may develop more of a bureaucratic career inclination than would be the case if they were employed by local authorities. Role-expectations and promotion possibilities seem dictated more by conformity to the system rather than by changing the system.

The training of teachers, the status-payment classification of teachers and the splitting of teachers into status groups by the type of school in which they work, all tend to create barriers in attempts to change teacher education. Teachers are known as Grundschullehrer, Hauptschullehrer, Realschullehrer, or Gymnasiumlehrer. This reduces communication between groups and encourages the desire of the higher status teachers, often in decision making positions, to work against change.

Finally, the fact that at least two distinct teacher associations exist makes overall change problematic.

The Training of Administrators

Special training for educational administrators is still not available in most German states. However, there is a growing recognition that the school of the future will require a trained administrator in charge (Mueller et al., 1973; Deutscher Bildungsrat, 1970, 1973).

The school of the future, constantly changing and adapting to needs, will have functions that are not apparent in non-changing schools. Some functions are supervision, planning, in-service education, evaluation, group leadership, motivation and stimulation. The administrator would need special education in those areas as well as in research techniques and organizational analysis. This opens up a whole new field in graduate education for practising administrators and for teachers who plan to move into administration.

Perhaps the clearest statement of reform in administrative practices is provided in a recent report by the Deutscher Bildungsrat (1973). The proposals follow logically from the 1970 proposals for restructuring educational systems with a view of achieving greater equality of educational opportunity and an education more in line with societal needs. The new school structures would call for new administrative structures.

Two basic themes support the proposals: an increased school autonomy and the participation of teachers, students, and parents in decision making. The two themes are not supposed to be seen as single innovations but as developmental processes which need to be initiated and gradually developed. They are to be incorporated together with changed administrative structures in schools.

Autonomy of Schools. The autonomy of schools would depend on the competence of the actors involved, and would operate within guidelines set by the Kultusministerium. It is made clear that the autonomy is only a limited one because the guidelines will establish the general aims and specific objectives, the organization of students for learning, basic methodological guidelines, materials and media, the number of courses and the marks required for graduation and minimum rules for curriculum development. In each case there will be considerable latitude so that teachers will be able to adapt their materials and methodology to the needs of the students.

The teachers will be the main decision makers in this framework. They will be formally organized in *Lehrerkonferenzen* (teacher conferences) where teachers are to be in the majority. The decision on the planning and organization of instruction in a school will be presented for critical comment to students and parents who will also be formally organized, but without formal authority.

Most personnel and financial matters will continue to be handled by the central authority even though school participation in these matters should be gradually expanded, so that schools could make significant decisions on problems peculiar to their own environment.

Participation of Teachers, Students, Parents. As already indicated, increased participation in school decision making is recommended for those directly affected by the school's activities.

The three groups would be formally organized, receive a small expense account, and have scheduled opportunities to examine decisions, and make recommendations. Committees can be created to work at specific problems or plans. However, students and parents will not be permitted to decide on any issues that relate to schools.

The Administration of Schools. The *Schulrat* (administration) is in a position of influence as well as of responsibility, both to the Kultusministerium for the interna as well as to the *Schultraeger* for the externa. He is subject to a dual leadership role. In the present organizational model his source of authority comes from above and usually he is mostly concerned with carrying out the directives that come down to him. Whether this new recommendation would change this is highly questionable, since his tasks still are largely outside the school. Little change is recommended in regards to *Schulraete*.

There are, however, certain recommendations that may provide a new orientation for the school administrator. First, it is recommended that the principal be elected by the teachers. This will give the administrator a third source of authority, that from below and he might generate more leverage to obtain concessions for his school. The term of service is to be for a limited time only. In larger schools, a team of administrators is recommended.

Clearly, the adoption of such changes could have a far reaching effect on educational programs in German schools. It would also be open to serious risks because of the absence of trained administrators, because of the strength of teacher groups, and because of the lack of attention to a series of organizational functions such as control, supervision, evaluation, leadership, and long range planning. All of these developments point to a high priority for the training of educational administrators for German schools.

Resistance to Change

Perhaps because of these risks no recommendations seem to appear on changing several crucial aspects of German education. Several long standing practices and organizational structures of German educational systems seem to continue unchallenged, even during discussion of significant change and innovation. Among those that stand out as significant to the observer are the following:

Administrative Structure. The dual authority in educational matters has not been challenged. The hierarchy of authority upward from the school principal, where the authority moves directly down from the Kultusminister, remains unquestioned. Even the roles of the incumbents of various positions in the hierarchy, together with the upward movement of aspiring individuals and the placement of legally trained people in decision making places is rarely challenged.

Early Selection of Students. Though the selection for the three streams of education is somewhat veiled by several mild innovations, it is still as much in force as always. The positive change is that more and more students find places in university entrance schools (Gymnasium). There is a great reliance on testing in the German educational scene. Even before the first grade, the youngster must "make the test" that psychologists have prepared before being admitted to grade one. Should he fail, then he is placed in a peculiar institution known as a Vorschule, or if below six years of age, is denied school entrance altogether. After four years, the child writes further tests, the results of which taken together with the Grundschule recommendation, determine which of the three school streams he or she pursues. To soften the blow or to allow for a margin of error, pupils may delay this experience for one or two years. The Gesamtschule (comprehensive school) was supposed to combine the three streams but in many cases it has simply brought the three streams into one building or onto one campus. Where genuine attempts are made to integrate the three streams, the teachers and administrators are worried lest the *Abitur* examinations will deal a severe blow to their efforts.

The Abitur. The graduation from Gymnasium is accomplished through the successful completion of examinations in carefully prescribed academic disciplines. This Abitur is the key to university or professional education.

Core Curriculum. The humanities, science, or language routes of the Gymnasium stand highest in importance in German education. Mathematics and languages still are the undisputed subjects for the best students.

The Individual Student. The student generally considers it a privilege to be able to enter Gymnasium. His needs, his desire for freedom and participation are submerged in the demand for conformity and acquisition of subject matter content and related skills. Rarely does one hear an argument for humanizing the classical schools.

Centralization. Financial, curriculum, personnel, policy, planning, and evaluation decisions are generally made at the highest level. Even ideas for innovation and change seem to emanate from the ministry of education. Local control of educational matters is not understood too well by the German people, and it will take time, effort and planning to get students, parents, teachers and administrators meaningfully involved in educational decision making.

Conclusion

A recognition of the need for change is present at the highest and lowest levels of the educational hierarchy in Hessen, West Germany. However, historically, change of the educational structures has been very slow in this country. There is little evidence that the reliance on the traditional system will dissipate, nor that the change, except in peripheral

matters, will accelerate until such time as the German people will turn to some of the fundamental issues in their quest for change. This, too, seems unlikely, except through a sudden catastrophic break, until the ethos of unquestioned acceptance of authority and tradition diminishes in the German mind. Again the school cycle reinforces this ethos in numerous ways making the process somewhat of a closed system. To create real change, system breaks related to traditional organization and practices will have to be deliberately engineered in the closed system as it now operates. In business, communication, industry and transportation, Germany adapts quickly and effectively. In education, it has yet to find its way to adapt to a new world view.

References

- Brock, A. *Lehrlings Handbuch*. Darmstadt: Metzler Verlag, 1972.
- Butts, R. F. *A cultural history of Western education*. New York: McGraw-Hill, 1955.
- Der Hessische Kultusminister. *Bildungswege in Hessen* (A series of pamphlets).
- Deutscher Bildungsrat. *Empfehlungen der Bildungskommission*. Stuttgart: Ernst Klett Verlag, 2 Auflage, 1970.
- Deutscher Bildungsrat. *Empfehlungen der Bildungskommission*. Teil I. Bonn: Bundesdruckerei, 1973.
- DGB Gewerkschaft. *Bildungspolitik*. Konferenzbericht. Frankfurt am Main, 1973.
- Edding, F. *Auf dem Wege zur Bildungsplanung*. Braunschweig: Westermann, 1970.
- Geisler, E. E. *Autoritaet und Freiheit*. Regensburg: Julius Klinkhardt, 1970.
- Gesamtschule Nordweststadt Frankfurt am Main*. Europaische Verlagsanstalt GMBH, 1968.
- Gewerkschaft Erziehung und Wissenschaft, Kampf gegen Genscher. Entwurf, Nummer 12, Dec. 1, 1973.
- Gewerkschaft Erziehung und Wissenschaft. *Die integrierende und differenzierende Gesamtschule*. Frankfurt an Main, (not dated, approx. 1970).
- Hauke, H. *Aspekte der Kuenftigen Schule*. Heidenheim an der Brenz: Heidenheimer Verlagsanstalt, 1970.
- Husen, T. International Project for the Evaluation of Educational Achievement (IEA): International study of achievement in mathematics. A comparison of twelve countries, Bd 11. Stockholm, 1967.
- Jensen, S. *Bildungsplanung als Systemtheorie*. Bielefeld: Bertelsmann Universitaetsverlag, 1970.
- Mueller, H., Vogelsang, H., & Bessoth, R. *Rationalisierung, Bildungsplanung und Management in der Schule*. Koeln: Verband Bildung und Erziehung, 1973.
- Preissler, G. *Raumordnung und Bildungsplanung*. Frankfurt am Main: Max Traeger Stiftung, 1968.
- Rahmenrichtlinien*. Frankfurt am Main: Verlag Moritz Diesterweg, 1972.
- Rolff, H. G. *Bildungsplanung als rollende Reform*. Frankfurt am Main: Diesterweg, 1970.
- Roehrs, H. *Das Gymnasium in Geschichte und Gegenwart*. Frankfurt am Main: Akademische Verlagsgesellschaft, 1969.
- Schultze, W., & Fuehr, C. *Schulwesen in der Bundersrepublik*. Weinheim: Beltz, 1966.
- Tanner, L. N., & Tanner, D. We must be doing something right. *Educational Leadership*, 1974, 31 (6), 567-568.
- The Times: Higher Education Supplement, October 19, 1973.
- von Hentig, H. *Systemzwang und Selbstbestimmung*. Stuttgart: Ernst Klett Verlag, 2 Auflage, 1970.
- von Hentig, H. *Was sollen Schueler lernen?* Frankfurt am Main: Fischer Taschenbuch Verlag, 1973.

MARJORIE S. ARIKADO

Ontario Institute for Studies in Education

Status Consensus as it Relates to Team Teacher Satisfaction

The relationship between the variable status consensus and team teacher satisfaction was examined using 134 teacher teams distributed among seventy-one open plan schools in Ontario. Two operational definitions of status consensus were tested. A significant relationship was found to exist between competence of a formal team leader as perceived by team members, and their satisfaction with the team teaching situation; however, the degree of team agreement on the highest ranking team member (in terms of leadership qualities) was not found to be significantly related to team satisfaction. These findings represent a part of a larger study which included additional independent variables. (Dr. Arikado is Assistant Professor of Educational Administration at the Ontario Institute for Studies in Education.)

Teacher satisfaction with the team teaching situation was selected as an area of study in an attempt to isolate some of the variables which relate to the apparent dissatisfaction with this type of teaching organizational structure (Musella & Selinger, 1972). In this study the teaching team is defined as "a group of teachers who share major responsibility for the instruction of the same group of students, and who coordinate their instructional activities among themselves" (Molnar, 1972, p. 1). Instructional activities are interpreted to include joint planning, teaching and evaluation of pupils, with much of the emphasis on planning and evaluation. The independent variables include team leadership, team size, and status consensus.

There are a number of reasons for the selection of status consensus as a variable worthy of study. In reviewing the area of small group studies, Heslin and Dunphy (1964) found that three variables accounted for a major proportion of variance associated with member satisfaction with the group. Status consensus was one of these variables. Of thirty-seven studies reviewed, thirteen had findings which "support the notion that status consensus is an important dimension underlying variations in member satisfaction" (p. 100).

By comparing interaction patterns for laboratory groups with high and low member satisfaction, Bales (1953) attempted to specify the conditions under which a preponderance of positive reactions over negatives ones is most likely to develop. "Such a state appears to be brought about most readily by the group's reaching status consensus, defined as consensus about the relative status of all group members" (Heslin & Dunphy, p. 100).

In the search for an operational definition of status consensus, Shelley (1960) suggested an alternative definition of status consensus which stressed consensus about the highest ranking group member. In testing this definition using non-laboratory groups, Shelley found that this measure of status consensus correlated more highly with group member satisfaction than did Bales' measure which relied on the ranking of all members by all group members.

Several laboratory studies have provided results which indicated a negative relationship between imposed leadership and status consensus. Cohen, Bennis and Wolkon (1962) found that the existence of an appointed leader created a lack of status consensus which was resolved by the establishment of an informal leader after a change-over to a new group formation, and that member satisfaction was significantly highest for those groups which replaced their appointed leader with a leader of their own selection. Similarly, Goldman, Bales and Martin (1961) found that groups without imposed leaders expressed "significantly greater enjoyment" than did those with imposed leaders. From these results, it appears that in organizational settings, the achievement of a high level of status consensus appears to be facilitated where group members acknowledge the leader's competence in relation to the group task (Heslin & Dunphy). In a study related to this, Hamblin, Miller and Wiggins (1961) found that perceived competence of the leader controlled a significant amount of the variance of morale.

In the study described in this paper, the relationship between status consensus and team satisfaction with the team teaching situation was examined. Two operational definitions of status consensus were used. The first one which was proposed by Shelley may be restated as the degree of team agreement on the assignment of a team member to the first (highest) rank. The second definition which was proposed by Heslin and Dunphy may be restated as follows: if the team has a formal leader, consensus is the degree to which the team rates the leader highly in terms of adequacy as a choice for the position. In this study, the formal leader was interpreted as that person who was formally selected by the principal to serve as leader of the teaching team.

Method

Sample

The following criteria governed the choice of schools selected for the study:

- 1) total staff size of not less than ten and not more than thirty-three teachers (although the number of team teachers could be fewer than this),
- 2) grades taught ranged from kindergarten to grade eight (K-8),
- 3) in operation for more than one year, and

4) open plan physical facilities.

Criterion 1 was designed to eliminate very small schools where family-type relationships are likely to predominate, and very large schools where bureaucratic operations and treatment of personnel are more likely to occur. Criteria 2 and 4 were included so as to control for relative similarity in teaching environment, while criterion 3 was included to eliminate schools going through the normal first-year period of adjustment.

The following criteria governed the choice of teachers selected for the study:

- 1) teachers belonged to teaching teams which had at least three full-time members (leaders are included in this count); and
- 2) teachers had operated as a team for at least six months prior to the collection of data.

Since small group theory points to dyads as having special characteristics which differ from those found in groups larger than two (Hare, 1962, ch. 8), criterion 1 eliminates pairs of team teachers, while criterion 2 eliminates newly-formed teams, thereby eliminating the factor of newness and problems associated with the initial stages of group development, both of which could contaminate the results.

Teams with formal leaders were identified by the school principal and the following criteria were imposed on the teams identified: team members must agree that their team does indeed possess a formal leader, and all members identify the same formal leader.

A total of 529 team teachers constituted the sample. These teachers belonged to 134 teams from 71 schools distributed among five county boards.

Instrumentation

Satisfaction with the team teaching situation, defined and operationalized in terms of favourable attitudes toward the team teaching situation, was measured by means of the semantic differential. (Osgood & Suci, 1955; Osgood, Suci, & Tannenbaum, 1957). In applying the semantic differential, each teacher was asked to rate his/her team teaching situation on a series of bipolar scales each of which was selected for this study because it had a high loading on the evaluative factor and low loadings on the other factors. Each of these scales consisted of a seven-point continuum with a negative adjective at one extreme and a positive adjective at the other.

For Heslin and Dunphy's definition of status consensus which incorporates the degree of team agreement on the rating of the formal leader, each team member was given five alternatives ranging from 1 representing an "excellent choice" for the position, to 5 representing a "poor choice" for the position. With respect to Shelley's definition of status consensus, i.e., the degree of team agreement on the highest ranking team member, team members were asked to select the person in their team whom they consider ranks highest in competence, ability to provide group guidance, influence over the group, and importance as a team member.

In order to obtain measures of team agreement on the highest ranking team member, an entropy formula was used. This formula, which was tested by Shelley (1960) and compared with other indices which measure degree of agreement, was found to have the highest correlation with

satisfaction. The formula was borrowed from entropy measures of thermodynamics and was found to provide an appropriate index of the degree of randomness in a distribution of rankings. Entropy, that is, the degree of randomness in the distribution of first rankings in the group, may be defined by the following formula:

$$H(I) = - \sum_i p(i) \log_2 p(i)$$

Where $p(i)$ is in the proportion of the first choices received by a given group member. This entropy value $H(I)$ equals zero if one member of the group receives all the first choices and is maximal when every group members selects a different first choice (the set of $p(i)$'s are then all equal). Shelley adopted this formula to be applied to groups of varying size. In so doing he explains that since the value of $H(I)$ increases with the number of group members, $H(I)$ may be divided by its maximal value for the size of the group. The maximal value of $H(I)$ is $\log_2 m$, where m is the number of group members. Thus the relative entropy formula

$$H_{\text{rel}}(I) = \frac{- \sum_i p(i) \log_2 p}{\log_2 m}$$

provides an index of randomness which can be used to compare groups of varying sizes (Shelley, 1960, pp. 159-160).

Results and Discussion

In using Heslin and Dunphy's operational definition of status consensus, the data from those teams having a formal leader were analyzed separately. The following hypothesis was tested using the correlation coefficient (r):

The degree to which the team rates the formal leader highly on adequacy as a choice for the position is positively related to the mean level of team satisfaction with the team teaching situation.

This hypothesis was confirmed; $r = 0.575$, $p \leq .01$ ($N = 34$). Confirmation of this hypothesis not only served as a test for Heslin and Dunphy's definition of status consensus (1964, p. 103), but it also provided support for the study conducted by Hamblin and his colleagues (1961) in which it was found that competence of the leader, as perceived by the group members, controls a significant amount of the variance of morale.

In using Shelley's definition of status consensus, all teams, both those with and without a formal leader ($N = 134$), were included and the following hypothesis was tested:

The degree of agreement in a team on first rankings is positively related to the mean level of team satisfaction with the team teaching situation.

Multiple linear regression analysis was used to test this hypothesis. Satisfaction with the team teaching situation was predicted on the basis of five categorical predictors which may be explained as follows: a team receives a 1 for one of five categories (which represent ranges of entropy values) of degree of team agreement on first rankings and zeros in the remaining categories; ranges of entropy included the values (a) $0 \leq x < 0.2$, (b) $0.2 \leq x < 0.4$, (c) $0.4 \leq x < 0.8$, and (e) $0.8 \leq x \leq 1.0$. The relationship

tested by this hypothesis was not found to be statistically significant at the .05 level.

Since this hypothesis was not statistically confirmed one might begin by questioning Shelley's conceptual model of status consensus and its applicability to the team teaching situation. Shelley's sample consisted of thirteen girls' clubs participating in the Y-teen program of the local YWCA. Looking at the size of the groups sampled (ten to thirty members to a club), the age of the group members (early teens), and the purpose of this organization, one may observe that these groups differed considerably from teacher teams. Also one may observe that the criteria for determining the highest ranking group member were different in each of the two studies. Shelley (1960) described the highest ranking club member as the one who was "most valuable to the club" (p. 161), whereas the study using teacher teams defined the highest ranking member as the one who ranks highest in terms of a combination of such qualities as competency, ability to provide group guidance, influence over the group, and importance as a team member. Many team teachers in this study expressed difficulty in selecting the highest ranking team member because they felt that all team members contributed more or less equally to the total team effort in a variety of ways. Some of those teachers on teams with formal leaders admitted that they ranked their leader as highest solely because of his/her formal position.

In conclusion, it appears that a re-examination of Shelley's operational definition of status consensus would be in order. There is need for clarification of the criteria used to determine the highest ranking group member. More important, as the results of the study using teacher teams indicate, Shelley's definition is not generalizable to all groups. Outcomes of any ranking procedure could well be influenced by the role and perceptions of persons doing this ranking, the criteria used to describe the highest ranking member, and the nature and purposes of the group doing this ranking.

References

- Bales, R. F. The equilibrium problem in small groups. In T. Parsons, R. F. Bales, & E. A. Shils (Eds.), *Working papers in the theory of action*. Glencoe, Ill.: Free Press, 1953.
- Cohen, A. M., Bennis, W. G., & Wolkon, G. H. The effects of changes in communication networks on the behaviors of problem-solving groups. *Sociometry*, 1962, 25 (2), 177-196.
- Goldman, M., Bolen M., & Martin, R. Some conditions under which groups operate and how this affects their performance. *Journal of Social Psychology*, 1961, 54, 47-56.
- Hamblin, R. L., Miller, K., & Wiggins, J. A. Group morale and competence of the leader. *Sociometry*, 1961, 24 (3), 295-311.
- Hare, A. P. *Handbook of small group research*. New York: Free Press of Glencoe, 1962.
- Heslin, R., & Dunphy, D. Three dimensions of member satisfaction in small groups. *Human Relations*, 1964, 17 (2), 99-112.
- Molnar, S. R. F. *Teachers in teams: Interaction, influence, and autonomy*. Stanford, California: Stanford Center for Research and Development in Teaching, 1972.

Status Consensus as it Relates to Team Teacher Satisfaction

- Musella, D. F., & Selinger, A. D. *Interim report on inservice education for open area schools*. Unpublished report. Toronto: The Ontario Institute for Studies in Education, 1972.
- Osgood, C. E., & Suci, G. J. Factor analysis of meaning. *Journal of Experimental Psychology*, 1955, *50* (5), 325-338.
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. *The measurement of meaning*. Urbana, Ill.: University of Illinois Press, 1957.
- Shelley, H. P. Status consensus, leadership, and satisfaction with the group. *Journal of Social Psychology*, 1960, *51*, 157-164.

W. PRITCHETT

Philadelphia Public Schools

and

D. J. WILLOWER

Pennsylvania State University

Student Perceptions of Teacher Pupil Control Behavior and Student Attitudes Toward High School

While there have been a large number of studies on teacher pupil control orientations and several on teacher pupil control ideology and student attitudes, the present investigation is the first dealing with teacher pupil control behavior and student attitudes.

A sample of 852 secondary school pupils responded to the Pupil Control Behavior Form and Coster's High School Students' Opinion Questionnaire. A significant relationship between student perceptions of custodial teacher pupil control behavior and negative attitudes toward school was found. This pertained to overall student attitudes and to each of six factors resulting from a factor analysis of Coster's instrument. (Dr. Pritchett is assistant director, Division of Music Education, Philadelphia Public Schools; Dr. Willower is Professor of Education, The Pennsylvania State University.)

Background

This article reports a study of teacher pupil control behavior and student attitudes toward school. It is one of a series of inquiries on the school as a social system and the place of pupil control in that system.

This line of work is based on a conceptual framework that depicts the public school as an organization characterized by a marked degree of political vulnerability and charged with the educational processing of a diverse but captive student clientele likely to hold views and values of its own. Hence, internal management and control receive high organizational priority, and various structures such as routines and rules serve as adaptive mechanisms that limit variety and reduce uncertainty for the organization and its personnel. In this kind of setting, teachers and students face

one another at cross-purposes, the former concerned with discipline and instruction, the latter with adjusting to and coping with the essentially adult system (Willower, 1971a, 1971b).

Pupil control is thus seen as a salient feature of school life. The topic has been addressed in more than 70 studies dealing with educator pupil control ideology (PCI) and behavior (PCB) (Willower, Eidell, & Hoy, 1973). Two instruments have been utilized as operational definitions for educator PCI and PCB, both based on a custodial-humanistic pupil control continuum. A custodial approach to pupil control stresses the maintenance of order and a rigid stance toward students, while a humanistic approach emphasizes a more flexible orientation based on the view that students can be self-disciplining and responsible (Willower *et al.*, 1973; Helsel & Willower, 1974).

Only a few previous investigations have examined the relationship between teacher pupil control orientation and student variables. In a study of the PCI of high school teachers and student alienation, it was found that the dimensions of alienation defined as sense of powerlessness, normlessness, and self-isolation were associated with custodialism in the PCI of school faculty (Hoy, 1973; Rafalides & Hoy, 1971). Another researcher reported that middle schools and junior high schools with custodial faculties had greater student absenteeism and suspension rates than schools having more humanistic faculties, but there were no significant differences between the school types with respect to student truancy and tardiness (McBride, 1972). In a third investigation, seven high schools that had experienced student unrest were matched with an equal number without student unrest. It was found that teachers in the schools characterized by student unrest were more custodial in PCI than their counterparts in the matched schools (Duggal, 1969).

These inquiries all dealt with teacher pupil control ideology. The research reported here is the first in this series of studies to examine teacher pupil control behavior in relation to student attitudes.

Hypothesis

The research just reviewed shows relatively negative student responses to be associated with custodialism in PCI. Further, the proposition that rigid teacher behavior will beget negative attitudes toward school on the part of students is a plausible one, and it should especially pertain in societies that espouse equalitarian values. Hence, the major hypothesis guiding this investigation was stated as follows: there will be a positive relationship between custodialism in teacher pupil control behavior and unfavorable student attitudes toward school.

Methods

Instruments

In order to test the major hypothesis, operational definitions for teacher PCB and student attitudes toward school were required. The PCB Form devised by Helsel and Willower was employed to measure the former. This instrument, which is completed by students, taps teacher PCB on the humanistic-custodial continuum already discussed. Responses to the 20-

item form are made on a five point scale ranging from "always" to "never." Examples of items, prefaced by the words "my teacher," are "says nice things to students," "is bossy with students," and "smiles when students are around." Eight of the items are stated in custodial terms and twelve are stated in a humanistic direction. A high score indicates a custodial PCB. Item-scale correlations for the PCB Form averaged .81, and a reliability analysis yielded a coefficient of .92 as estimated by Cronbach's alpha (Helsel & Willower, 1974).

The device used to measure students' attitudes toward school was Coster's High School Students' Opinion Questionnaire, part one, "How I Feel About My High School," hereafter called the HSSOQ (Coster, 1954, 1958). The 27 items that comprised this measure were scored using a five point set of response categories that indicated attitudes from very favorable to very unfavorable. Examples of items are "What, in general, is your opinion of the teachers in your high school?" "Do you feel that your high school work is the kind of work that you like to do?" and "Do you feel that going to high school will help you in enjoying life more?"

With respect to the reliability of the instrument, Coster reported that, for a sample of 878 high school students, "the coefficient of correlation between the sub-total scores made on odd- and even-numbered items was computed, using the Pearson Product Moment formula, and a value of .84 was obtained" (Coster, 1954, p. 38). The overall score on the HSSOQ is used to indicate students' attitudes toward school; the higher the score, the more unfavorable the attitude.

Coster indicated that HSSOQ items could be divided into seven groupings. These were attitudes toward teachers, the school, school program, appropriateness of school work, future expectations, social acceptance, and miscellaneous attitudes. Presumably, items were grouped on the basis of apparent similarities in meaning. Although Coster's groupings made sense in this regard, a factor analysis of the HSSOQ was performed as part of the present research.

Sample

These instruments were administered in the spring of 1973 to selected students in nine secondary schools located in a metropolitan area that included parts of two states in the northeastern United States. Four of the schools were junior high schools and five were senior high schools. Two of the junior high schools and three of the senior high school had predominately black student bodies, while the remaining four schools were predominately white. Predominate racial character was based on the criterion that at least 80 per cent of the student body be of one race.

The instruments were administered to approximately 90 students in each school. Depending on class size, students in from three to six classes in each school completed the forms during regular sessions. A total of 852 usable student responses were secured in this manner. The student, rather than the class, was the unit of analysis used in testing the major hypothesis and in examining the association of the various factors of the HSSOQ with PCB. Statistical treatments, therefore, are based on an *N* of 852.

Findings

The major hypothesis proposed a positive relationship between custodialism in teacher PCB and unfavorable student attitudes toward school. In order to test this hypothesis, a Pearson Product Moment Correlation technique was employed. The correlation between student statements of teacher PCB and student attitudes toward school was .31. This correlation was moderate, accounting for slightly less than 10 per cent of the variance, but it was significant beyond the .001 level. Hence, the major hypothesis cannot be rejected.

The factor analysis of the HSSOQ revealed significant patterns of loadings for six factors. (Factor loadings of .40 or above were considered significant) (Pritchett, 1974). Although five of the factors were named in accordance with Coster's original schema, there were a number of discrepancies in the items which constituted the original categories and their corresponding factors. The six factors of the HSSOQ are shown in Table 1 along with the correlation coefficients of each of these relatively independent subscales with the PCB scores. All correlations were significant at the .001 level. However, some of the correlations are quite low and account for a limited amount of variance. That the highest correlation obtained was for the factor concerned with attitudes toward teachers is not surprising.

TABLE 1
HSSOQ FACTORS AND CORRELATION WITH PCB SCORES

Factor	r
Attitudes toward:	
1. Appropriateness of school work	.23
2. School program	.12
3. Social acceptance	.14
4. School	.18
5. Social Context	.13
6. Teachers	.35

When PCB scores are examined by school (see Table 2), mean scores for the junior and senior high schools in the sample are quite similar. However, the mean PCB score for one of the junior high schools, school six, is comparatively humanistic, nearly seven points lower than that of the next most humanistic junior high school. If the score for school six is excluded, the mean for the remaining junior high schools, 52.1, contrasts more sharply with the mean for the senior high schools.

The mean PCB score for the predominately black schools is somewhat lower or more humanistic than the mean for the white schools. This finding is similar to that of another study in which comparable data were gathered (Brown, Willower, & Lynch, 1974).

TABLE 2
PUPIL CONTROL BEHAVIOR BY SCHOOL

School	Level	Racial Character	PCB Mean Score
1	JHS	W	52.4
2	SHS	W	49.1
3	SHS	W	47.0
4	JHS	W	54.0
5	SHS	B	51.5
6	JHS	B	43.3
7	SHS	B	47.5
8	SHS	B	47.7
9	JHS	B	50.1
<hr/>			
Means:	JHS = 49.9 SHS = 48.7	B = 48.0 W = 50.6	All students = 49.1

Although reliability estimates for the PCB Form were reported to be good, an additional reliability check was conducted using the present sample of 852 respondents. An alpha index of reliability of .91 was obtained, quite close to the .92 reliability reported in the original PCB work.

Discussion

We begin our discussion with a disclaimer. The correlation coefficients between the variables examined were low to moderate, even though they were significant. Moreover, no effort was made to control for the influence of other variables. Hence, our findings should be interpreted with caution.

Our main empirical results indicate an association between student perceptions of custodial teacher behavior and negative attitudes toward school and especially toward teachers. The relationship between perceived teacher pupil control behavior and student attitudes is quite direct and consistent. It holds for the overall attitude measured by the HSSOQ and for each of its factors, although it is less strong for certain dimensions.

It is not clear whether a direct cause-effect relationship pertains. However, it can be speculated that a snowball effect occurs. An earlier paper in which the term external rather than custodial control was used suggested that "... the use of external controls creates a snowball effect, in that school is made less attractive for pupils, increasing their alienation; and, in consequence, a still greater stress is placed by teachers on external controls, and so on" (Willower, 1965, p. 43).

If this happens, a kind of vicious circle in the affective domain may develop that could be difficult to break. Whether custodial teacher behavior would result in lowered student achievement in cognitive areas is problematic but plausible. In any case, this is a question open to inquiry.

The extent to which students' reports of teacher PCB coincide with "reality" is, of course, not known. The student was taken as the unit of analysis in the present study, so perceptions were given full sway. Since the dependent variable we wished to examine was student attitudes toward school, it made theoretical sense to focus on student perceptions of teacher PCB as the independent variable. The classic dictum that perceptions con-

stitute reality for the perceiver pertains. However, it should be stressed that research on teacher PCB ordinarily employs the classroom as the unit of analysis. When students tend to agree in their descriptions of teacher PCB, confidence is added that their collective perceptions reflect something more real than the perceptions of given individuals. The perception-reality gap cannot be bridged with finality, but unless there are compelling theoretical reasons to the contrary, investigations of teacher PCB should be based upon collective descriptions of behavior. Since the current inquiry was not based on group descriptions of PCB, it is essentially a study of perceptions and no claim concerning the reality of teacher PCB is made.

A final point deals with the issue of student preferences for teacher PCB. While we found that negative student attitudes are associated with perceptions of custodial teacher PCB, the matter of student preferences was not directly addressed. This question, which also can lead to analyses of the relationships of student characteristics and preferences, is being examined in an exploratory study currently underway. This and other work should help to shed more light on the nature and consequences of pupil control behavior in schools.

References

- Brown, L. H., Willower, D. J., & Lynch, P. D. School socioeconomic status and teacher pupil control behavior. *Urban Education*, 1974, 9 (October).
- Coster, J. K. Factors related to morale in secondary school. Unpublished doctoral dissertation, Yale University, 1954.
- Coster, J. K. Attitudes toward school of high school pupils from three income levels. *Journal of Educational Psychology*, 1958, 49.
- Duggal, S. P. Relationship between student unrest, student participation in school management, and dogmatism and pupil control ideology of staff in the high schools. Unpublished doctoral dissertation, University of Michigan, 1969.
- Rafalides, M., & Hoy, W. K. Student sense of alienation and pupil control orientation of high schools. *High School Journal*, 1971, 55 (December).
- Helsel, A. R., & Willower, D. J. Toward definition and measurement of pupil control behavior. *Journal of Educational Administration*, 1974, 12 (May).
- Hoy, W. K. Dimensions of student alienation and characteristics of public high schools. *Interchange*, 1973, 3.
- McBride, A. P. A comparative study of a group of New Jersey middle schools and junior high schools in relation to their pupil control ideology and selected pupil behaviors. Unpublished doctoral dissertation, Rutgers University, 1972.
- Pritchett, W. The relationship between teacher pupil control behavior and student attitudes toward school. Unpublished doctoral dissertation, The Pennsylvania State University, 1974.
- Willower, D. J. Hypotheses on the school as a social system. *Educational Administration Quarterly*, 1965, 1 (Autumn).
- Willower, D. J. Social control in schools. In L. C. Deighton (Ed.), *The encyclopedia of education*. New York: The Macmillan Co. and The Free Press, 1971.(a)
- Willower, D. J. The teacher subculture. In L. W. Drabick (Ed.), *Interpreting education: A sociological approach*. New York: Appleton-Century-Crofts, Inc., 1971.(b)
- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology*. University Park: Penn State Studies No. 24, 2nd ed. with an annotated bibliography of pupil control studies, 1973.

C. A. McCATTY

University of Toronto

Patterns of Learning Projects Among Professional Men

The purpose of this study was to investigate the extent and some of the characteristics of adult learning. Subjects were a random sample of 54 professional men.

Data were collected by an intensive interview. Findings support the view that adults engage extensively in learning. Vocational subject matter formed the largest category. Self-planned learning dominated the field of participation. (Dr. McCatty is Assistant Professor in the School of Physical and Health Education, University of Toronto.)

Although adults have always continued to learn throughout life, there are many factors which suggest the possibility of an increased incidence of deliberate learning among adults. One major motivation has been the necessity of re-learning job skills.

There are many other motivating factors which stimulate the adult to participate in learning. Increasing amount of leisure, current emphasis on group effort, and the demands of operating a democratic society have placed learning requirements on adults.

Research indicates that men and women are participating extensively in learning efforts. Johnstone and Rivera (1965), in a national survey, discovered that, in the United States, in one year, just over 20 per cent of adult Americans were actively participating in some educational activity. Later research studies (e.g., Blackburn, 1968) have suggested the existence of even more deliberate adult learning efforts than had been revealed in Johnstone and Rivera's study.

In spite of the extensive participation in continuous learning and the endeavours of many researchers in the area, there are still numerous lacunae in the knowledge of the extent and nature of adult participation. Before 1971, researchers employed interview techniques which elicited only that learning which the interviewee could recall fairly quickly and easily. Thus, many gaps remained in the knowledge and comprehension of adult learning participation.

Tough (1971) introduced an intensive, probing type of interview which appeared to be remarkably successful in uncovering a large number of learning efforts. The present study employed the same type of interview.

Problem

This study focussed on questions about learning of professional men.

To what extent do adults participate in learning?

What subject matter do adults learn?

What proportion of adult learning is motivated by the desire for credit and what proportion is motivated by considerations other than credit?

How much adult participation in learning occurs: (1) in a group, (2) with an individual on a one-to-one basis, (3) with a non-human resource (such as a programmed instruction book or a set of tape recordings), (4) when the learner is directing his own learning, and (5) in a mixture of two or more of the previous sources with no one source predominant? What proportion of group learning participation by adults is planned by a group directed by an instructor and what proportion by a peer group? What are some characteristics of the peer groups? In the learner-directed projects, what is the adult's primary method of learning? (Does he read or discuss or observe or what in order to learn?) In mixed projects, what percentage of the learning is assigned to each category of learner-teacher relationship and what is the chronological order of participation among categories?

One type of reason for distinguishing between different types of learner-teacher relationship was the next area in which further research was indicated. Why do adults choose the type of learner-teacher relationship they do for planning their learning? What relationship is there between subject matter and choice of planner?

Review of the Literature

Many researchers have considered the extent of learning participation by adults (Armstrong, 1971; Blackburn, 1968; Brown, 1964; Ingham, 1964; Johnstone & Rivera, 1965; Litchfield, 1965; Tough, 1971). Depending on the procedures and definitions used, there were varied findings by these researchers as to the percentage of adults who had recently conducted a learning project. The highest incidence was that reported by Tough; when he interviewed a group of 66 adults from seven different occupations, he found that 98 per cent had participated in a deliberate, sustained effort to learn in the previous 12 months.

The content of adult learning was considered by Blackburn (1967) and Johnstone and Rivera (1965). These researchers found that adults study a wide variety of subject matter.

The proportion of credit and non-credit learning has received attention from some researchers (Johnstone & Rivera, 1965; Tough, 1967). Johnstone and Rivera concluded that credit was not an important motivation for learning by adults.

Johnstone and Rivera (1965) were the first researchers to explore the extent of independent study being conducted by adults; they found that nearly 8 per cent of adults had engaged in an independent study in the previous 12 months. None of the earlier studies found as much self-planned learning as did the Tough survey (1971); two-thirds of the learning projects con-

ducted by Tough's interviewees had been planned by the learner himself.

Some studies have served to give some indication of resources employed by self-directed learners. Tough (1967) concluded that his 40 subjects obtained a large amount of human assistance in their self-teaching projects. Johnstone and Rivera (1965) identified recordings and tapes as aids which their subjects said they would employ if they taught themselves. Shorey (1969) listed the publications which Windsor teachers used in their continuing education efforts.

Many researchers have explored various factors which influence an adult's method preference (Blackburn, 1967; Butterdahl & Verner, 1965; Jenkins, 1960; Johnstone & Rivera, 1965; Koenig & McKeachie, 1959). Some of the factors investigated were age, occupation, income, education, and social and personality characteristics. No research could be found where the learner himself was asked to express his reasons for choosing a type of planner for his learning.

Johnstone and Rivera (1965) investigated the relationship between subject matter and methods of study and concluded that such a relationship existed. Blackburn (1967) explored method orientation and reached the conclusion that part of the variation in adult method orientation was due to the nature of the subject matter.

No other study could be found which scrutinized the learning conducted by a whole spectrum of professional men. The only studies referred to in this section which employed a probing interview were those by Armstrong (1971) and Tough (1971).

Method

Population

The population in this study consisted of males between the ages of 30 and 50, with at least one year of post-secondary education, with a minimum annual income of \$12,000.00, and in occupations rated high on the socioeconomic scale. The scale employed for identifying the occupations was Blishen's scale (1968).

The Setting

The population was limited to those who were resident in the Borough of North York. The Borough of North York is one of the six member municipalities of Metropolitan Toronto. It has an area of 69.5 square miles and a population of approximately 500,000.

One major advantage of this area for this study was that there are numerous choices of methods for educational participation. North York possesses Canada's third largest educational system; there are 109 elementary schools, 25 junior high schools, 17 secondary schools, and 2 vocational schools. The Borough Board of Education has a Department of Adult Education. This educational system offers a wide variety of adult educational programmes—academic, arts and crafts, languages, technical and commercial. York University's two campuses are situated in North York; through its Centre for Continuing Education, the University provides courses for adults in many areas. The North York YMCA specializes in adult education; numerous courses are made available for adults through

the YMCA. Seventeen community centres operate in the Borough, offering a choice of several educational opportunities for adults.

Sample

The sample was selected from the assessment rolls of the Borough of North York, prepared in 1970. These rolls consist of a series of large books, in which sections are assigned to each of the 12 wards. The pages assigned to each ward are numbered consecutively, starting with page 1.

The researcher selected the ward number and page number by employment of a table of random numbers. He then started at the top of the selected page and proceeded in order down the list, scrutinizing 45 address slots and noting the occupations of the individuals. Where a male's occupation coincided with one of those being considered in this study, a record was made of his name and address. To obtain the required number of interviewees, the researcher had to make 105 such draws.

A subsequent letter and telephone call sought the subject's cooperation and checked the criteria of occupation, age, and income. The acceptance rate of those who were requested to participate in the study was 90 per cent. The 54 interviewees consisted of 14 engineers, 9 lawyers, 9 teachers, 8 doctors, 4 dentists, 3 architects, 2 professors, 1 chemist, 1 geologist, 1 judge, 1 pharmacist, and 1 physicist.

The Learning Project

Learning participation was considered here in terms of learning projects. The learning project was the central phenomenon of this study. Essential to the concept of the learning project, as employed here, is a minimum time of seven hours during which the primary purpose is to learn. The seven hours must be compressed into a six-month period (thus eliminating a desultory type of effort). The seven-hour (or longer) project is normally composed of a series of learning episodes (although, theoretically, it could consist of one long episode). Also included may be some episodes devoted to planning, preparing, or travelling, connected with the learning, but, after the time spent on these episodes has been excluded, the total participation must be not less than five hours.

The learning project is concerned only with sustained and highly deliberate efforts. The main purpose of these efforts is to learn, i.e., to effect a change in the individual himself. The learner must be clear as to what he wants to learn. These are the three criteria which are the very core of the definition.

There is no restriction on the type of change sought by the learner. It may be acquisition of new knowledge or skill, change of personality, or improvement in physical functioning. The changes which occur in the learner may be easily observable, as they would be in learning to ski. Or the changes may be the subtle alterations in maximum oxygen uptake, haemoglobin count, and pulse rate which accompany an improvement in physical fitness.

The definition of a learning project specifies that the learner must want to retain the knowledge and skill which he gains. The length of time for retention was set arbitrarily at two days.

Data Collection

The data used in this study were collected by means of an intensive interview. The model for the interview schedule was Tough's schedule (1971), which was employed in 1970 to investigate similar learning phenomena. Prior to use in this study, the interview schedule was revised several times on the basis of trial interviews conducted with each version of the schedule on the investigator's friends.

A major objective of the schedule was to familiarize the interviewee with the concept of a learning project. Another very important objective in the formulation of the interview schedule was to devise a very probing technique, which would elicit as many of the subject's learning projects as possible.¹

Most of the interviews took place in the subject's office or home. During the process of the interview, repeated oral probes were made to uncover as many learning projects as possible which the interviewee had conducted in the previous 12 months. When these probes failed to uncover any more learning projects, two handout sheets were employed to aid the subject to remember his learning projects. One of these sheets listed some of the things that adults learn, and the other listed some of the ways that adults learn.

The subject was asked to estimate the number of hours that he spent on each learning project in the 12-month period prior to the interview.

The interviewee was asked if any part of his total motivation for undertaking any of the learning projects had been the desire for credit. If more than 50 per cent of the total motivation had been the desire to obtain a degree, diploma, or certificate, or for an examination or upgrading related to a job, the learning project was categorized as a credit project.

The remainder of the interview discussed the planner of each learning project. (The planner was here defined as the person, group, or object which made the day-to-day decisions regarding what and how the subject learned.) The subject was asked first to identify the planner for each learning project.

The subject was then asked to categorize the planner in one of four categories—group, one-to-one, non-human, or self (described on a page of the schedule which was handed to the interviewee). In *group* learning, the learner is a member of a group of five or more persons (including the instructor); the group may have a specific instructor or may be a group of peers. In *one-to-one* learning, there is one learner (or sometimes two or three) and one instructor—a common type of relationship between the learner and the golf professional, the driving school instructor, or the piano teacher. In *non-human resource* or *inanimate* learning, an inanimate resource (e.g., a programmed instruction book or set of tape recordings) plans the learning project. In *learner-planned* or *self-planned* learning, the learner assumes the primary responsibility for planning and conducting the learning project. A project was categorized as *mixed* if the responsibility for planning did not reside primarily (51 per cent) in any one of the four above-defined categories.

If the planner was in the peer group category, the subject was asked to describe the group briefly. If the planner was in the learner-planned

category, he was asked to name his major resource. For projects in the mixed category, the subject was requested to assign the percentage of planning to each category and to name the order in which the two or more planner categories operated.

After all learning projects had been categorized, the subject was asked regarding one of the non-credit learning projects in each of the four pure categories why he had chosen the planner which he did. (If more than one learning project was in the category, one was selected by use of a table of random numbers.)

Findings

Participation

The participation of the 54 learners was considered, first, in terms of the number of projects, and then in terms of the number of hours.

The 54 subjects had participated in a total of 599 projects in a 12-month period. This was an average of 11.1 projects for each subject; the median was 10.3. No interviewee had failed to participate in any learning project at all. The least number of projects participated in by an interviewee was 2, and the greatest number was 31. The mode was 11, and 7 subjects had participated in that number of projects. This represented a considerable extent of participation, as far as number of projects was concerned.

TABLE 1
HOURS OF PARTICIPATION

Number of Hours	Number of Subjects	Number of Hours	Number of Subjects
0- 99	0	1700-1799	2
100- 199	1	1800-1899	1
200- 299	2	1900-1999	1
300- 399	2	2000-2099	1
400- 499	2	2100-2199	0
500- 599	3	2200-2299	1
600- 699	5	2300-2399	0
700- 799	4	2400-2499	0
800- 899	6	2500-2599	0
900- 999	1	2600-2699	1
1000-1099	3	2700-2799	1
1100-1199	3	2800-2899	0
1200-1299	2	2900-2999	1
1300-1399	4		
1400-1499	1		
1500-1599	2		
1600-1699	2	4200-4299	2

The full story of the extent of participation is not told, however, until the measurement by time is also considered. Data in Table 1 show the number of hours of participation for the 54 subjects over a 12-month period.

None of the subjects had participated for fewer than 100 hours, and only one for fewer than 200; he had spent 157 hours in learning projects. Two of the subjects participated about 4,200 hours.² The total number of hours of

participation for 54 subjects was 67,200, or an average of 1,244 hours per subject; the median was 1,058 hours.

In general, then, it was found that the interviewees had participated extensively, whether judged in terms of number of projects or in hours of participation. The findings revealed a much higher incidence of participation than had been discovered by earlier researchers (Blackburn, 1967, 1968; Johnstone & Rivera, 1965).

Subject Matter

This study employed Johnstone and Rivera's subject matter categories (1965) for classification of the 599 projects. The seven categories of subject matter identified in this study will now be discussed.

Vocational. By far the largest segment of subject matter was in the vocational category. More than 50 per cent of all learning projects for the sample in this study were work-oriented.

A fairly common type of vocational learning project was keeping current in the interviewee's vocational area. This was a continuous type of effort; the project had no end.

Other vocational learning projects had a limited time span and were undertaken in response to a specific problem. One example of this type was the project conducted by a lawyer who represented a client in a motor accident and had to learn the details of the accident and the previous application of the law in similar cases. A dentist's motivation for a learning project was a new discovery in his profession—an implanting technique.

Hobbies and recreation. The hobbies and recreation category contained 89 projects and was the second most extensive category. The projects in this group, by definition, were carried on in leisure time, and had no work connotations.

The greatest number of projects in a sub-category (27) were associated with outdoor recreation, and included skiing, sailing, golf, skating, flying, nature, snowmobiling, riding, driving, camping, conservation, fishing, and navigation. A doctor and an architect had chosen both skiing and sailing as learning projects, and another doctor had learned to sail and to ride.

Sports (spectator), music, art, technical arts and hobbies, travel and other countries, and agriculture were some of the other classes of subject matter proposed as learning projects in the hobbies and recreation category. In technical arts and hobbies, photography, with 4 projects, was the most frequently reported learning effort. Travel and other countries (7 projects) was a fairly sizable sub-category. The subjects were not arm chair travellers; in 6 of the 7, there was actual travel related to the project, and in the seventh, travel was contemplated.

Current events. The current events category contained 56 projects, and 39 learners had participated in these projects. The subjects included here were those dealing with current social, political, and economic affairs and civic responsibilities.

Most of the projects were aimed at having a general grasp of current topics, but in many projects, a specific area of interest was identified. Among these areas were the Israeli Six Day War, Ontario politics, and minority struggles. Five interviewees learned about pollution, and one

wrote a brief on the subject for the environmental department of the government.

Home and family life. In the home and family life category, there were 56 projects. These projects included topics pertaining to the establishment, maintenance, and improvement of a home, or to the carrying out of household duties and family responsibilities.

In the area of home improvement skills, there were 19 projects. There were 17 home economics projects, 9 projects in gardening and landscaping, 3 in family relationships, and 1 in future family prospects. One interviewee had learned how to cook.

Personal development. There were 28 projects in the field of personal development. This category was defined as being comprised of subjects aimed at helping people expand themselves in the areas of physical fitness, health, personality development, and interpersonal and social skills. Nearly all (26) of the projects were in the area of physical fitness. Of the remaining 2 projects, one was sensory awareness, and the other was conscious relaxation. No respondents had undertaken a learning project aimed at improving their personality or social skills.

Academic and general education. In the academic and general education category, 25 projects were reported. These were academic subjects normally studied as part of a high school or college education and excluded all job-related subjects.

The projects reported by the interviewees included history, literature, anthropology, philosophy, political science, psychology, and languages. A geologist's interest in anthropology had been aroused as a result of attendance at a seminar in this area sponsored by Trinity College for its alumni. An engineer's learning project in Dickens' works began because, being Dutch, he had never read Dickens, and his wife gave him some of Dickens' books for Christmas. A doctor learned Italian because he had some Italian patients.

Religion, morals, and ethics. The smallest group of subject matter was in the religion, morals, and ethics area; there were 15 projects. Included were projects concerned with traditional religious training, religion applied to everyday life, and all other aspects of religion, morals, or ethics. An engineer's longest project (435 hours) was in the religious area when he and his family became born-again Christians.

Ten Types of Learning Projects

Learning projects were categorized into types by a cross-tabulation of two dimensions. One dimension dichotomized learning, on the basis of motivation, as credit and non-credit. The other dimension identified four distinct types of learning on the basis of the planner of the learning—group, one-to-one, inanimate, and learner-planned. When the majority of the planning of a learning project resided with one of these four planners, the project was assigned to that category. In this planner dimension, a fifth type of situation was recognized and termed mixed learning; the planning was conducted by a mixture of planners from two or more categories, with no one planner responsible for more than 50 per cent of the planning.

The interviewee was asked to specify the planner of each project and to assign the planner to one of the four categories. Where a project was

planned by planners from more than one category, it was assigned to the planning category in which the majority of the planning lay (providing more than 50 per cent of the planning was conducted by a planner in any one category). If no planner was more than 50 per cent responsible, the project was classified as mixed.

This fragmentation of the learning participation into 10 different cells permitted a discussion of how the learner participates along the two dimensions which form the 10 types of learning projects identified in this study. Is the learner primarily motivated by the desire for credit? Does he choose as a planner a group, an individual, a non-human resource, or himself?

TABLE 2
PROJECTS AND PARTICIPATIONS IN 10 CELLS

Planner	Credit				Non-Credit				Total Projects	
	Projects		Participants		Projects		Participants			
	No.	%	No.	%	No.	%	No.	%	No.	%
Learner	1	0	1	2	457	76	54	100	458	76
Group	7	1	3	6	62	10	34	63	69	11
One-to-one	1	0	1	2	40	7	24	44	41	7
Inanimate	0	0	0	0	4	1	4	7	4	1
Mixed	0	0	0	0	27	5	17	31	27	5
All Five Categories	9	1	3	6	590	99	54	100	599	100

Table 2 provides a comparison of learning participation in the 10 cells. The extent of participation in each cell is shown (measured by the number and percentage of projects, and number and percentage of participants). The percentages shown in Table 2 indicate the incidence of participation in each cell measured in terms of projects (599 = 100 per cent) and participants (54 = 100 per cent).

The last line of Table 2 is really a comparison of the credit and non-credit participation. These figures disregard the planner distinctions and report the results for the total credit learning and total non-credit learning.

The comparison indicates that 99 per cent of all projects were non-credit, and only 3 participants reported credit projects. The total hours of credit participation for the 54 subjects was 657, and the total hours of non-credit participation was 66,543.

Of the 9 credit projects, 7 were group-planned. A single individual accounted for the 2 credit projects which were planned in the one-to-one and learner-planned cells (as well as 4 of the 7 projects in the group-planned cell).

Another surprising fact indicated by the data in Table 2 is the extent to which learner-planned learning dominated the field of participation. All in-

interviewees had participated in self-planned learning, and 76 per cent of the projects were in this category. The range of projects per participant in this cell was 1 to 28, and the mean and median were 8.5 and 7.5 respectively.

The inanimate planner was used to a very limited extent by the interviewees; only 4 participants had employed this type of learning.

Of the 34 participants in the non-credit group cell, 20 had one project, and the 3 who had participated most extensively had 4, 5, and 6 projects respectively.

In both the non-credit one-to-one cell and the non-credit mixed cell, the range of projects per participant was 1 to 4.

In 27 of the projects, the interviewee could not allocate the major planning responsibility to a planner in any of the 4 categories recognized in this study. Because no single category of planner was responsible for more than 50 per cent of the planning responsibility, these 27 projects were listed as mixed planner, in accordance with the definition of mixed planner. In most cases, the planning was equally divided by planners in two categories; in many cases, the subject and his wife planned together on an equal basis. These mixed projects were all non-credit-motivated. There was a fairly substantial volume of mixed planning. In extent, the mixed planner learning was just below the 3 main categories (group, one-to-one, and learner-planned). Of the 27 projects, 23 shifted between planners in no definite chronological order. In many cases, the shuttling back and forth was so rapid that the planning became "meshed".

Only 20 per cent of the interviewees had used only one planner category (in each of the 11 instances, the category used being the learner-planned). Most of the subjects had used two or three different categories of planner.

Of the 62 non-credit group projects, 10 (or 16 per cent) were conducted in a peer group and 52 with an instructor. In all 7 of the credit group projects, the group had an instructor. Of the 10 peer groups, 7 were committees which had been appointed, or elected, to carry out one phase of an entire operation, and the committee members had to learn certain things in order to fulfil their function.

A total of 25 interviewees reported having at least one project in the one-to-one area (24 of them with non-credit motivation). The number of projects reported was 41.

Methods for Learner-planned Learning

For each learner-planned project, the learner identified the source of the majority of his subject matter. In 391 of the 458 projects, one major method was identified by the interviewee. In 59 cases, 2 methods were combined equally, and in 8 cases, there was an equal combination of 3 methods.

The most common single method was reading. A variety of reading materials was employed, and sometimes identified specifically by the interviewee. Books, newspapers, journals, magazines, and reports were all credited as sources of information.

The second most common method for the learner-planned projects was discussion—sometimes with one person, and sometimes with several.

Doing was identified in 35 projects as the major method of learning.

This method might be employed to provide a trial and error type of learning situation, or to effect a change in the subject through practice.

Observing was a less common method of learning. Only 2.8 per cent of the total 458 projects depended on observing as the major single method of learning. There were, however, 15 other projects in which observing was combined equally with another method, and 3 in which observation was combined with 2 other methods, with each responsible for one-third of the total.

Reasons for Choice of a Planner

Articulated reasons. Other researchers have considered various personal, social, and personality factors which are judged to be influential in the adult's choice of a planning category. In this study, it was decided to identify the conscious reasons, expressed by the participants, for choosing the planning category which they did. From each planner category reported by each interviewee, one project was chosen at random, and the interviewee was asked why he had chosen the planner he had specified for that project. In some cases, more than one reason was given in answer to this question; all specified reasons were recorded.

The 54 participants who reported self-planned projects gave 61 reasons for self-planned participation.

The most frequently cited basis for choice was desire for individualized subject matter; 25 participants gave this as a motive. Thirteen of the 25 wanted to select the subject matter so that it could be applied to their particular interests. The other 12 had very specific problems; their learning consisted of acquiring the knowledge to solve these problems.

Physical availability was cited in 10 instances as a reason for choice of a learner-planned approach. It was an economy of time or effort for the learner to do his learning in the most convenient location for himself.

In 7 cases, the learners stated that a certain facet of their personality made them choose a self-planned type of learning. A feeling of independence and self-reliance (or, perhaps, in some cases, lack of self-confidence) prodded these learners towards an autonomous approach to planning.

Five of the interviewees identified their expertise in the subject matter as their reason for planning the projects themselves. The projects were in their professional area.

Financial economy accounted for four decisions by learners to plan for themselves.

Flexibility of time was a factor inducing the learner to plan his own project in four cases; a course would have imposed the requirement of specified hours of attendance.

In three cases, the learner was at the stage in the subject matter where he thought that self-planning was more appropriate than other types. Two of these learners considered that they were sufficiently advanced that they did not require outside help in the planning. The third participant said that he had not advanced sufficiently to warrant joining a group.

The effectiveness of the method, the ease of the subject, and his own

capability at planning were each given once as a reason for using the self-planned approach.

A total of 37 reasons were expressed for choice of the group-planned category. The most frequent of these (9) was that that type of format already existed in a form such as a staff or committee meeting arrangement. The interviewee therefore participated in the existing manner.

The capability of the instructor was given 8 times as the reason for a group type of learning and was by far the most frequent reason expressed for choosing the one-to-one type.

Subject matter. In this section, the relationship of subject matter to planner choice is discussed. The large majority of projects in the vocational and hobbies and recreation categories were learner-planned. No other area of subject matter had such a large percentage of learner-planned projects as current events (96 per cent). Of all the subject matter areas, none had as large a percentage of one-to-one projects as did the personal development area; 29 per cent of these projects were planned on a one-to-one basis. The largest percentage of group-planned projects for any subject matter area was found in the religion area, where 47 per cent of the projects were planned in a group setting.

Conclusions and Implications

The data support the major premise in this study regarding the extent of learning by adults, and especially of the learner-planned type. The participation by adults is extraordinarily extensive. The adults interviewed in this study spent an average of 24 hours each week participating in learning projects of different types, including 20 hours in projects which the learner had planned for himself. This did not include the casual learning which is part of every human's life. Also excluded was purposeful learning in which a series of related learning episodes were of shorter duration than 5 hours.

The data also indicate that when the adult undertakes self-planned learning (100 per cent of the interviewees did so), he does so mainly because he can tailor the material he learns to suit his individual needs.

Taken together, these findings have certain implications for the practitioners of education—both in the more obvious scholastic institutions (schools, colleges, and universities), as well as in those professions which provide continuous professional education. Since such a major part of the adult's learning is planned by himself, it is apparent that a primary aim in education should be to equip the adult with skill in planning learning. The present emphasis on what the student should learn might be profitably supplanted with an emphasis on how the student should learn.

The findings regarding the extent of learning also have implications for researchers in this field. The type of interview employed in this study, which used some probing questions and two hand-out sheets, appears to be very effective in aiding in the recall of learning efforts.

When the learners planned their own learning, they relied heavily on reading and observation for provision of the subject matter. Any improvement in the learner's acuity in these areas would aid in his ability to learn

and might also form part of a general body of information on learning skills and techniques.

The capability of the instructor was a very important factor in the learner's choice of methods. In fact, the instructor's proficiency may be decisive in the initial decision to learn. The implication of this fact for programme planning is important; the provision of expertise may trigger off the desire to learn.

There are implications in this research for practitioners of physical education. Keeping physically fit was a learning project for 26 of the 54 interviewees. The majority of their physical fitness projects were learner-planned, and no projects were performed in a group setting. The practitioners of physical education should augment their present group-oriented programmes with campaigns aimed at providing the most recent knowledge in physical fitness to the individuals who will be planning their own programmes.

The subjects in this study engaged in a large body of learning as a leisure pursuit. Yet, learning has been largely overlooked by researchers as a leisure activity. It should be recognized that learning projects can constitute a satisfying leisure pursuit.

One finding in this study was that, in very few learning projects, was credit a major motivation. This knowledge is of importance to many planners of adult learning (e.g., extension departments of universities). Course offerings, for some segments of the public, need not provide the incentive of credit, in the form of a degree, diploma, or certificate, in order to be attractive to most learners.

One purpose in the selection of the population used in this study was the expectation that their learning might act as a model for others. The sheer immensity of their learning effort alone should stimulate others to more extensive participation. A psychological barrier often seems to inhibit human accomplishment until there is an awareness of others' successful endeavours (a well-known phenomenon in athletics).

This study has shown that adults can plan a great deal of learning for themselves. An awareness of the potential which we all possess to teach ourselves may inchoate many learning projects which might have remained latent because of difficulties which the learner experiences in seeking or obtaining a teacher.

This article is based on a Ph.D. thesis, Ontario Institute for Studies in Education, University of Toronto, 1973. The thesis is obtainable on microfilm from the National Library of Canada, Ottawa K1A 0N4.

¹ To indicate the technique used, the following sections have been extracted from the interview schedule and are reproduced here.

Is there anything you have spent several hours in learning over the past 12 months?

Now think back over the past 12 months. Have you made any deliberate effort to learn to do anything, or to obtain information, or increase your knowledge? Have you tried to increase your understanding or change your emotional reactions?

I am interested in any effort you made to learn anything at all. It doesn't have to be important; it doesn't have to be hard or serious; it can be easy, or trivial, or fun. We can include any information, knowledge, or skill, or understanding that you have tried to gain, just as long as you spent a few hours at it. Can you recall anything else?

We're trying to get as complete a list as possible, so think back now right to (month) of last year. It doesn't matter what the knowledge or skill or information was, just as long as you spent several hours at it. What else can you recall?

² The top participant in terms of hours (4,233 hours) was a dentist. His participation was so extraordinarily extensive that he was queried subsequent to the interview. He stated that he considered his estimate a good one.

Patterns of Learning Projects among Professional Men

He attributed the extent of his participation to two factors. Firstly, his current events project included 500 hours he spent listening to news broadcasts on the radio while at work. Secondly, he slept only four to five hours nightly, leaving himself more leisure time than most people; most of his leisure was devoted to learning. Like Houle's learning-oriented individual (1963), learning with him was a constant activity.

References

- Armstrong, D. P. Adult learners of low educational attainment: The self-concepts, backgrounds, and educative behavior of average- and high-learning adults of low educational attainment. Unpublished doctoral dissertation, University of Toronto, 1971.
- Blackburn, D. J. *Method orientations of adults for participation in educative activities*. (Doctoral dissertation, University of Wisconsin), 1967. Dale County Title Co. Microfilms.
- Blackburn, D. J. *Guelph adult participation patterns*. Guelph: Department of Extension Education, University of Guelph, 1968.
- Blishen, B. R. A socio-economic index for occupations in Canada. In B. R. Blishen, F. E. Jones, K. D. Naegle, & J. Porter (Eds.), *Canadian society: Sociological perspectives*. Toronto: MacMillan of Canada, 1968.
- Brown, A. M. College experience and continuing educational activity. In D. Solomon (Ed.), *The continuing learner*. Chicago: Center for the Study of Liberal Education for Adults, 1964.
- Butterdahl, K., & Verner, C. Characteristics of participants in two methods of adult education. *Adult Education*, 1965, 15 (2), 67-73.
- Houle, C. O. *The inquiring mind*. Madison: The University of Wisconsin Press, 1963.
- Ingham, R. The relationship of educative behavior to the leisure satisfaction of college alumnae. In D. Solomon (Ed.), *The continuing learner*. Chicago: Center for the Study of Liberal Education for Adults, 1964.
- Jenkins, D. H. Conditions underlying good learning. In M. V. Miller (Ed.), *On teaching adults: An anthology*. Chicago: Center for the Study of Liberal Education for Adults, 1960.
- Johnstone, J. W. C., & Rivera, R. J. *Volunteers for learning: A study of the educational pursuits of American adults*. Chicago: Aldine, 1965.
- Koenig, K., & McKeachie, W. J. Personality and independent study. *Journal of Educational Psychology*, 1959, 50, 132-134.
- Litchfield, A. *The nature and pattern of participation in adult education activities*. (Doctoral dissertation, University of Chicago) Chicago: Department of Photoduplication, University of Chicago Library, 1965. No. T 12244.
- Shorey, L. L. Teacher participation in continuing education activities. Unpublished doctoral dissertation, University of Toronto, 1969.
- Tough, A. M. *Learning without a teacher: A study of tasks and assistance during adult self-teaching projects*. Toronto: The Ontario Institute for Studies in Education, 1967.
- Tough, A. M. *The adult's learning projects: A fresh approach to theory and practice in adult learning*. Toronto: The Ontario Institute for Studies in Education, 1971.

JOAN C. BREITER

Iowa State University

Reading or Listening: A Comparison of Techniques of Instruction in Elementary Social Studies

The relative value of reading and listening as techniques of instruction to facilitate comprehension of social studies material was explored. Sixth grade students, grouped by intelligence level, reading ability, and sex read or listened to identical material drawn from commercial supplementary units and were tested on their comprehension of this material. Resulting data indicated that the comprehension of children of high intelligence, high and below average reading ability, and girls was significantly better when reading than when listening on at least one of the two units, though neither technique facilitated a significant difference in the comprehension of all sixth grade children. (Dr. Breiter is Assistant Professor of Elementary Education, Iowa State University.)

The teacher who seeks information concerning the relative merits of learning by reading and of learning by listening may have difficulty identifying guidelines which clearly indicate the advantages of one method over the other.

Both reading and listening are considered key activities in learning factual material. Listening is the original source of information for children, but has the disadvantage that "seldom can it be mulled over or even accurately recalled by the listener" (McKee, 1966, p. 34). Reading skills, however, later become important learning tools in every area of study, as noted by McKee (1966) when he stated, "From the third grade on, reading can and should provide advantages that listening cannot supply" (p. 13). In general, research has suggested that people do learn more by *listening* during the early school years, but increasingly learn more by *reading* during the later elementary school years. Many reasons might be suggested for this conclusion, chief among them possibly being that while children are normally exposed to in-depth, sequential instruction in reading skills, much less instruction is given in listening skills. Until fairly recently, not enough has really been known about listening, the various listening skills,

and how to teach them to give effective instruction in this area. Largely, it has been assumed by instructors that children just knew how to listen.

This study does not attempt to specifically advocate either of these skills—reading or listening—as the ultimate in instruction. It is common knowledge that some children learn better by one mode than by another, and many prefer or even require exposure to information by more than one approach. It does, however, attempt to look at the present status of reading and listening as techniques of instruction within the context of one grade level in one subject area, and to assess which seems to be most likely to be effective with which students. In this context, it is hoped this study will assist the classroom teacher in gaining some insight into choosing appropriate teaching techniques for individuals and groups in the classroom.

Related Research

Research in the fields of reading and listening has been voluminous and varied. A brief description of several studies which have a pertinent bearing on the purpose of this paper follows.

Listening. Maurice S. Lewis (1953) attempted to measure the ability of intermediate grade pupils to comprehend some of the materials commonly presented orally in the classroom with two equivalent forms of a test which measured with some degree of accuracy the following five listening skills: (1) ability to get main ideas, (2) ability to recall facts and details, (3) ability to make inferences from facts presented in a selection, (4) ability to recognize word meanings from context, and (5) ability to follow directions (p. 1). Lewis assumed that listening ability, like other human abilities, is distributed among the general population in accordance with the principle of normal probability, that individuals differ in their ability to comprehend the spoken word, and that listening comprehension can be measured in quantitative terms (p. 2). His results seemed to imply that the level of reading difficulty has little or no bearing upon the level of listening ability, and that there is a great difference in the listening comprehension of individuals (p. 33).

Ann Jackson's (1966) study proposed to determine if a difference exists in the listening comprehension of boys and girls, and to determine the correlation between listening and various reading abilities. She concluded there was a significant correlation between listening and reading vocabulary and also between listening and reading comprehension. She found no significant difference in the ability of boys and girls in listening comprehension (p. 85).

A study in listening comprehension by Charles A. Lindsey (1953) reported that pupils' listening comprehension appeared to be more dependent upon the interest appeal of the material being presented than upon such factors as classroom interruptions, etc. (p. 98). Though not a surprising finding, since most people have found themselves "shutting out the world" when highly involved and interested in something to which they are listening (or reading), it is a most interesting finding in regard to this project.

Composite Points of Interest—Listening. Overall, a survey of available literature in the field of listening yields the following points of interest which have been documented by one or more research studies.

1. Children are beginning to gain more by reading than by listening as they enter the upper elementary grades.
2. Some of the claims made for the superiority of listening comprehension may have been overrated if the material used in the studies was of extremely high interest to the pupils.
3. Listening may be a superior technique of instruction for some individuals, but is probably not so for all.
4. There are specific factors which affect the ability of children to comprehend through listening, such as the speed at which the material is presented, and the individual's adjustment.
5. Listening tests have been constructed which measure skills which researchers feel are indicative of the degree of listening ability possessed by an individual.
6. Ability to listen varies with the individual just as other skills.

Reading. Studies which explore various facets of reading comprehension in social studies are also numerous. Much of this work, such as a study by Lattin (1952), indicates that a great deal of children's factual knowledge of social studies material is less than adequate; less, in fact, than what authorities in the field considered to be common knowledge of the topic (Canada) over which the children were tested (p. 21). A partial explanation for findings such as this may be found in George Millis' (1959) study of pupils' understanding of social studies terms. In his work with fifth grade students, Millis found that 61.3 percent of their responses to designated terms drawn from their texts were vague and indefinite, or completely incorrect and/or not understood (p. 198). Similar-type results were found in a study by Walter E. Kerr (1934, p. 129).

In an investigation of text materials, R. H. Franzen (1922) found that much of the material was much too difficult for the sixth graders who were expected to use them (pp. 83-85). Inasmuch as Fred A. Sloan's (1959) study supported this finding, by noting that the general trend of publishers was to go above the indicated grade level (p. 206), there appear to be certain difficulties inherent in reading in the social studies which must be taken into account when choosing reading as a technique of instruction.

Composite Points of Interest—Reading. Overall, when considering the available literature in the field of reading in the social studies, the following generalizations, documented by one or more research studies, may be made.

1. Reading, especially when limited to textbooks, is not a completely adequate technique of instruction.
2. The use of textbooks as the sole source of reference in social studies is inadequate.
3. The criticism that textbooks are often too vague and underdeveloped in their treatment of concepts and generalizations appears to be at least partly valid.
4. In reading materials of various kinds and of various levels of difficulty, it is necessary to adjust the rate at which the material is read. Specific instruction in this skill may help to improve a child's reading comprehension.
5. If reading comprehension is to occur, use of reading materials of an appropriate level is necessary.

6. The chance of success in learning social studies generalizations is enhanced by the use of a variety of methods and materials, among them reading and listening, but also including other aids and approaches.

A study of double interest is Richard Hampleman's comparison of the listening and reading comprehension ability of fourth and sixth grade pupils. Some of Hampleman's findings will be supported by this study, some will be opposed, and some will suggest related factors. Because of its application to the study being reported, Hampleman's conclusions will be reported in the discussion of results later in this paper.

Procedure

This study proposed to determine the relative value of two specific techniques of instruction in the social studies, reading and listening, when the factors of varying levels of intelligence and reading ability, and sex were considered. The general purpose of the study was to determine if there was a significant difference in the comprehension of those students who read the chosen material and those who listened to it. The following eight categories of students were considered within the context of the general question, "What is the effect of reading and the effect of listening on the comprehension of supplementary social studies materials by sixth graders?"

1. students of high intelligence
2. students of average intelligence
3. students of below average intelligence
4. students of high reading ability
5. students of average reading ability
6. students of below average reading ability
7. male students
8. female students.

The procedural steps of this investigation were as follows. Two supplementary social studies unit texts, each dealing with a single topic, were chosen as the experimental material. These unit texts were commercial material written for Grosset and Dunlap, Inc. as part of their Who-When-Where series of supplementary social studies materials for upper elementary grades. These texts were titled *Aztecs Of Mexico* and *This Is Canada*.

After the materials were chosen, all major concepts, facts, and generalizations in the unit texts were listed in the order in which they appeared. For each concept, fact, or generalization, a question or questions which reflected the basic idea was constructed in multiple-choice format.

Both the unit texts and the tentative questions constructed from them were submitted to a panel of authorities. Since the material concentrated on the social sciences of history and geography, the panel of authorities who judged the materials and the questions drawn from them consisted of professors of history, geography, education, and psychology at Mankato State College, Mankato, Minnesota. This panel made suggestions for improvement of the quality of the questions in content and form, and judged the quality of the unit texts as suitable for the purpose of the study.

Each unit text was divided into four lessons—consisting of roughly the amount that could be read or listened to in 20 to 25 minutes. The ap-

appropriate questions from the total number approved by the panel were organized into question sets—one for each lesson, and an answer sheet was devised which could be used for all question sets and for the pre- and post-tests also devised at this time. A pre-test and a post-test were constructed for each of the two unit texts. They were identical, except that the order of the questions in the post-test was a reversal of the order of the questions in the pre-test.

At this time a manual to be used by the teacher in administering the experimental materials was written. It included all needed information about the materials, the question sets and tests, and directions to the teacher for each lesson for each of the two unit texts. Directions for handling and returning the materials and the children's completed answer sheets were also included.

Considerable care was taken to ensure content validity for these experimental, investigator-written materials used with the unit texts. In detail, from the list of item ideas reflecting the facts, concepts, and generalizations found in the unit texts by chapters, one or more multiple-choice questions, each having four possible answers, were developed for each basic fact, concept, or generalization. These questions were grouped into sets which covered the lessons in detail. Four question sets were constructed in this way for each of the two unit texts. These question sets each included the concepts in one or two chapters of material, depending on the length of the chapters, the number of items involved, and the estimated time of reading or listening to the text material.

The pre-test and post-test questions were designed to test knowledge of facts, concepts, and generalizations more general in nature and of greater importance than many of the items in the question sets. The reversal of question order comprised an effort to minimize the possibility of memorization of test questions by the students.

The manuals consisted of directions as to procedure; purpose of materials; seating, timing, and other testing considerations; specific directions for the administration of each test; and directions for handling the materials. This material was meant to serve the following purposes: The pre-test for each unit text helped the investigator to determine the prior knowledge of the unit text topic each child possessed. The post-test helped to determine the knowledge of each child at the end of each unit. The question sets were basically a motivational device to encourage the children to listen or read carefully. The manual of directions was to assist teachers in maintaining a uniform presentation and administration of the materials in regard to time limits, approaches used, and amount of teacher explanation and assistance which could be given.

The submission of the tests, question sets, and manuals to experts in history, geography, professional education, and test construction for suggestions, criticism, and approval provided the basis for the content validity claimed for this material.

The experimental materials and unit texts were first used in two pilot studies with 111 sixth-grade students who comprised three classes. The first pilot study was made with two classes (87 pupils) and the second with one class (24 pupils). After each of these pilot studies, the materials were revised as indicated by item difficulty and discrimination analyses of the

items, and suggestions by the classroom teachers who used the manuals and materials.

The revised experimental materials were then organized into their final form and packaged into class-sized sets along with the appropriate unitext or set of tapes to be used. A description of the use of these experimental materials follows.

The directors of elementary education in several cities in southern Minnesota were contacted for the purpose of explaining the study and requesting permission to conduct it in their systems. An effort was made to identify systems which were considered, statewide, to have good educational programs and teaching staffs, and whose enrollment included a cross-section of pupils who would be both typical of the geographical areas and of varying reading ability and intelligence levels. Two systems were chosen to participate in the study. One was a college community (Community A) which had some degree of industry; the second (Community B) was more heavily industrialized. Both communities fell into the 25,000 to 50,000 population range.

With the permission of the respective director of elementary education of each system, all sixth grade teachers in these systems were invited to participate in the study. Fourteen teachers in seven schools in Community A, and nine teachers in five schools in Community B accepted this invitation. They then attended orientation sessions which prepared them for their role in the project, and familiarized them with the experimental materials. These teachers were responsible for twenty-eight classes of which fourteen were from each community.

The twenty-eight participating classes were then randomly assigned to the experimental (listening) group or the control (reading) group. In order that complete data would be available at the end of the study for each child in the sample population, the following limitations were made concerning the use of student scores in the analysis of data. A student was eliminated from the sample population if:

1. he was absent during any part of the period during which his class was participating in the project,
2. he had an excessive hearing loss, thus preventing him from using materials which were listening-based,
3. his tested intelligence was below that of children usually assigned to the "normal" classroom, or
4. his records did not contain a fifth-grade test score on either the *California Short Form Test of Mental Maturity* or the *Lorge-Thorndike Intelligence Test* and a sixth-grade silent reading test score on the *Iowa Test of Basic Skills* administered in the fall of that school year.

A schedule was set up which allowed each of the twenty-eight classes of sixth graders, under the direction of their usual classroom teacher, to use the experimental materials one hour per day for ten school days. Five periods were used to read or listen to the text of "Unit I" and complete the questions sets which matched each lesson. Five were similarly used for "Unit II". Each "unit" was preceded by a pre-test and concluded with a post-test. At the end of the experimental period, which was from November through January, complete data was available for 570 students. Of the 741

students originally participating, 171 were eliminated on one or more of the limitations given above; 286 comprised the experimental (listening) group, and 284 comprised the control (reading) group.

As previously indicated, a standard measure of intelligence and one of silent reading ability was recorded for each student participating as a member of the sample population. Though the sample population was basically divided into two major groups—reading and listening—three intelligence level subgroups which cut across these lines were established through the use of these test scores. The limits, in terms of intelligence quotients, of these subgroups were as follows:

Groups	Intelligence Quotient
Above average	110 and above
Average	90 - 109
Below average	below 90

Similarly, three reading ability subgroups which cut across the major grouping lines were established on the basis of standard scores. The limits of these subgroups were as follows:

Groups	Limits
Above average	one-half grade or more above grade level
Average	one-half grade below to one-half grade above grade level
Below average	more than one-half grade below grade level

The composition of the sample population by intelligence groups, reading groups, and sex was as described in Table 1.

TABLE 1
COMPOSITION OF GROUPS

Groups	Total	Control Group	Experimental Group
<u>I.Q.</u>			
Above Average	339	169	170
Average	206	103	103
Below Average	25	12	13
	570	284	286
<u>Reading</u>			
Above Average	348	178	170
Average	109	56	53
Below Average	113	50	63
	570	284	286
<u>Sex</u>			
Male	296	148	148
Female	274	136	138
	570	284	286

Treatment of the Data

At the conclusion of the experimental period, the data for the 570 students who made up the sample population was recorded. This data consisted of the pre-test and post-test scores of each child for the two units of experimental material, the child's IQ and silent reading comprehension score from the indicated standard tests, the sex of the child, and the intelligence and reading subgroups into which each child fell.

The level of significance established for this study was at the .05 level and analysis of covariance was used to analyze each group. In all cases the experimental (listening) group was contrasted with the control (reading) group, with the pre-test score used as covariate. The following analyses of covariance were made for each of the two units (unitexts).

1. Using only the students with above average IQ
2. Using only the students with average IQ
3. Using only the students with below average IQ
4. Using only the students with above average reading ability
5. Using only the students with average reading ability
6. Using only the students with below average reading ability
7. Using only the male students
8. Using only the female students.

Findings and Discussion

For each of the analyses relating to intelligence level, reading ability, and sex an F Statistic was established which had to be equalled or exceeded for a significant difference to exist due to instructional method. The majority of relationships showed no significant difference in the effectiveness of reading and listening as techniques of instruction. Areas in which a significant difference was found were as follows:

1. Children of above average intelligence learn more by reading than by listening.
2. Children of high and of below average reading ability *may* learn more by reading, as the statistics favored reading as a method of instruction for one of the two units for each of these groups. However, since this was not true for both units for either of these groups, the finding is not clearcut.
3. Girls learn more by reading than by listening.

The reader may at this time be interested in the conclusions of Hampleman's study, mentioned earlier. He found listening comprehension significantly superior for both sixth and fourth grade pupils—boys and girls—and, as might be expected, that sixth grade pupils were significantly superior to fourth grade pupils in both listening and reading comprehension. He also found easy materials to be more readily comprehended than hard materials by all groups in his sample population, and that listening comprehension showed greater superiority over reading comprehension with easy material than with hard material for both age groups. When considering the boys and girls as separate groups, he found boys superior to girls in the comprehension of hard material, but no apparent difference in the ability to comprehend long and short paragraphs. Similarly, the relationship between listening and reading comprehension did not appear to be altered by the length of the passage. Most significantly, in

relationship to the study being reported in this paper, an increase in mental age, and to a lesser extent chronological age, decreased the difference between listening and reading comprehension (Hampleman, 1955, pp. 83-84). Though Hampleman found evidence for listening as the more effective learning/teaching technique, this study found some support for reading. This finding may help to explain the difference in the conclusions reached. On the basis of the study being reported, it seems reasonable to conclude that while all sixth graders do not comprehend better by reading than by listening, certain groups have completed this transition, notably the high intelligence group and the female portion of the sample population. While there was no consistent significant difference in the adjusted means in favor of the reading group in the other situations, there was no group in which there was a significant difference in favor of the listening group. This may suggest that a majority of sixth grade children have reached the point in their development where they read in order to learn considerably more than they did in the primary grades, but not to the extent that reading has become their major avenue of learning. Listening remains an important method of learning for them, and will probably remain so throughout their lives, but it is no longer *the* major way for them to learn.

There are several limitations to this study. First, the listening portion of the experimental materials consisted of tape recordings of the written material read at an appropriate rate and with the type of expression one would use in reading a storybook to children. It was believed to be a valid use of the material, but it is clear that while materials in oral and written form may be comparable, they cannot be exact duplicates without biasing the study in the direction of written language. However, in the two areas in which reading seemed to have a significant effect, the use of written material for listening purposes may not have been a serious deficit. If the remaining results are interpreted in this light, it is possible that it would be found that listening would be significantly effective for some groups and reading for others—much as is found by the classroom teacher in her informal categorization of learning styles. A repetition of this study with the oral material cast into the form of a paraphrase of the written material would be of value in supporting or revising the findings of this study.

Second, the reader should consider the fact that children have received much formal reading instruction throughout the elementary grades, but are unlikely to have received much formal listening instruction. Also, the listening ability of these children was not measured. All data regarding the individual children concerning such items was restricted to what could be obtained from their cumulative records. No record of any listening test scores was included in these records, and the investigator did not receive permission to administer a listening test to these students. Undoubtedly, such information would have added much of value to the findings of this study. It is hoped that any replication of this study would include the listening skills factor as, combined with the factors of intelligence, reading ability, and sex, this variable would allow many additional interesting comparisons to be made.

Overall, the study presents several considerations to the elementary classroom teacher. Since reading presently does seem to be a more effective learning tool for at least a defined portion of sixth grade children, teachers

should make rather extensive use of it in their social studies instruction. The high intelligence children and the girls in the classroom are more likely to profit from this method of instruction, and may prefer it. Children with other levels of intelligence and reading ability may also use reading effectively—especially if care is taken to provide needed instruction as indicated above—but the teacher will need to include alternate instructional methods and materials, accompanied by appropriate direct instruction in the skills needed to use them effectively.

Since this study showed no significant difference in the effectiveness of reading and listening at this level for most of the groups represented in the study, attention should be given to providing a solid program of instruction in listening skills. Children originally gained most of their knowledge in this fashion, and should be given the opportunity to consciously increase their listening skills. Careful, direct instruction on specific listening skills, similar to what is routinely supplied in reading instruction, is of great importance to the individual who spends his life in a world largely tuned to listening both during his school years and throughout his life. Though much of his formal education will be through printed material, a significant portion of it will be presented orally; the working world of a large segment of the population rests only lightly on the written word. Children should not be penalized by being denied instruction in a mode of learning which will offer them another way to learn effectively. In recent years increasing attention has been given to developing useful tests of listening skills and instructional programs. A planned program of listening instruction should exist in every school.

There is much in this study which raises more questions than it answers. It should be replicated upon a basis which recognizes and takes into account the factors which have operated as a deficit in this study. Additional factors which might be considered in replication are:

1. conducting the study in areas of high or low educational opportunity and/or in areas of high or low economic status,
2. drawing the sample population from more than two school systems and from a variety of geographical areas, and
3. administering intelligence, silent reading, and listening tests to the sample population immediately prior to conducting the experimental part of the study.

The most important generalization the classroom teacher may draw from this study is that there are methods of instruction which are especially effective with specific children, and that the teacher can identify these children and the technique of instruction likely to be most effective with them. It is evident that more than one method of instruction should be used in the classroom, and that a combination of methods of instruction with such aids as pictures, films, television, etc., will likely prove to be more effective than any one technique used exclusively with all students. However, knowledge of what technique to use with which child is of great value when planning the educational experiences of children. With the increasing complexity of today's education, all information which adds to the teacher's insight in this area is of benefit to her and to the students for which she has educational responsibility.

References

- Franzen, R. H. Applying the criteria of comprehension to geography texts for elementary schools. Unpublished doctoral dissertation, Warwick and York, Baltimore, 1922.
- Hampleman, R. S. Comparison of listening and reading comprehension ability of fourth and sixth grade pupils. Unpublished doctoral dissertation, Indiana University, June, 1955.
- Jackson, Ann Elizabeth. An investigation of the relationship between listening and selected variables in grades four, five, and six. Unpublished doctoral dissertation, Arizona State University, 1966.
- Kerr, W. E. A case history of reading comprehension difficulties in geography material. Unpublished Master's thesis, State University of Iowa, 1934.
- Lattin, R. T. An evaluation of elementary school pupils' knowledge of Canada as related to the opinion of authorities. Unpublished doctoral dissertation, State University of Iowa, 1952.
- Lewis, M. S. The construction of a diagnostic test of listening comprehension of fourth, fifth, and sixth grade pupils. Unpublished doctoral dissertation, Colorado State College, 1953.
- Lindsey, C. The effect of certain factors on the listening comprehension of fourth, fifth, and sixth grade pupils. Unpublished doctoral dissertation, Colorado State College, 1953.
- McKee, P. *Reading: A program of instruction for the elementary school*. Boston: Houghton Mifflin Co., 1966.
- Millis, G. H. Fifth grade pupils' understanding of terms encountered in their social studies text. Unpublished doctoral dissertation, University of Illinois, 1959.
- Sloan, F. A., Jr. Readability of social studies textbooks for grades four, five, and six as measured by the *Dale-Chall Formula*. Unpublished doctoral dissertation, George Peabody College for Teachers, 1959.

List of Books Received

- Buros, O. K. (Ed.) *Test in print II*. Highland Park, N.J.: Gryphon Press, 1974, 1107 pp., \$70.00.
- Canadian Education Association. *A case for the school board research unit*. Toronto, 1974, 18 pp., \$2.00.
- Canadian Education Association. *Directory of education studies in Canada, 1973-1974*. Toronto, 1974, 164 pp., \$6.00.
- Crunican, P. *Priests and politicians: Manitoba schools and the election of 1896*. Toronto: University of Toronto Press, 1974, 369 pp., \$17.50.
- Crutchfield, M. *Individualized reading: A guide for teaching word analysis skills*. Los Angeles: Gramercy Press, 1975, 130 pp., \$4.95.
- Cushenbery, D. C., & Howell, H. *Reading and the gifted child*. Springfield, Ill.: Charles C. Thomas, 1974, 181 pp., \$9.50.
- DeCarlo, J. E., & Madon, C. A. *Innovations in education for the seventies: Selected readings*. New York: Behavioral Publications, 1973, 276 pp., \$9.95 (hardbound), \$4.95 (pb).
- Dunn, L. M. (Ed.) *Exceptional children in the schools* (2nd ed.). Toronto: Holt, Rinehart and Winston, 1973, 594 pp., \$11.50.
- Featherstone, J. *British primary schools today: An introduction*. Toronto: Macmillan, 1971, 62 pp., \$1.25.
- Fisher, H. (Ed.) *Developments in high school psychology*. New York: Behavioral Publications, 1974, 292 pp., \$12.95.
- Flood Page, C. *Student evaluation of teaching: The American experience*. London: Society for Research into Higher Education, 1974, 98 pp., U.K. price £2.50.
- Fritz, J. O. *My encounters with alternatives*. Toronto: Canadian Education Association, 1975, 36 pp., \$2.00.
- Levine, E. A. *Lisa and her soundless world*. New York: Behavioral Publications, 1974, 40 pp., \$4.95.
- Lupul, M. R. *The Roman Catholic church and the Northwest school question*. Toronto, University of Toronto Press, 1974, 292 pp., \$15.00.
- McCandless, B. R., & Evans, E. D. *Children and youth: Psychosocial development*. Toronto: Holt, Rinehart and Winston, 1973, 529 pp., \$11.50.
- Meyers, E. S., Ball, H. H., & Crutchfield, M. *The kindergarten teachers handbook*. Los Angeles: Gramercy Press, 1973, 103 pp., \$3.75.
- Piele, P. K., & Hall, J. S. *Budgets, bonds, and ballots*. Lexington, Mass.: D. C. Heath and Co., 1973, 224 pp., \$14.50.
- Rossi, J. J., & Filstead, W. J. *The therapeutic community*. New York: Behavioral Publications, 1973, 344 pp., \$16.95.
- Rowen, B. *The children we see*. Toronto: Holt, Rinehart and Winston, 1973, 316 pp., \$11.55.
- Sarason, I. G., & Sarason, B. R. *Constructive classroom behavior: A teacher's guide to modeling and role-playing techniques*. New York: Behavioral Publications, 1974, 52 pp., \$2.95.
- Schwartz, B. *Educating Man for the 21st century, Volume 8, Permanent education*. The Hague: Martinus Nijhoff, 1974, 246 pp., Netherlands price 37.50 guilders.
- White, M. A., & Duker, J. *Education: A conceptual and empirical approach*. Toronto: Holt, Rinehart and Winston, 1973, 359 pp., \$7.90.
- Wolfensberger, W. *The principle of normalization in human services*. Toronto: National Institute on Mental Retardation, 1972, 238 pp.

Research In Education

Edited by F. MUSGROVE and H. J. BUTCHER

CONTENTS OF NUMBER 13, MAY 1975

Articles on

The impact of social science on flows from school to university

Objective test analysis: some criteria for item selection

The satisfaction of university staff and their teaching

Syllabus-bound and syllabus-free orientations in a technological university

Some thoughts on norm-referenced and criterion-referenced measures

Book Reviews

RESEARCH IN EDUCATION is published twice-yearly in May and November, by Manchester University Press. Annual subscriptions are \$14.00 (post free); single copies are \$7.50

MANCHESTER UNIVERSITY PRESS
Oxford Road,
Manchester M13 9PL, England



The Alberta Journal of Educational Research

Volume XXI, No. 3

September, 1975

CONTENTS

A Logically First Stage in Training Creative Teachers?	143
<i>H. E. Poole</i>	
The Development of Word Meaning Discrimination in Children	154
<i>T. O. Maguire, R. B. Patsula, and P. O. Evanechko</i>	
Community Vocational Centres in Northwestern Alberta	168
<i>B. Y. Card</i>	
Performance of Canadian Native, Black and White Children on Some Cognitive and Personality Tests	183
<i>J. P. Das, J. Manos, R. N. Kanungo</i>	
Women Administrators and Women Teachers: A Comparative Study	196
<i>M. Nixon and L. Gue</i>	
Demonstrating the Relationship Between Values and Attitudes as a Means of Changing Attitudes	207
<i>J. W. Kehoe</i>	
The Effect of Noise and Object on Acquisition of a Sight Vocabulary in Kindergarten Children	213
<i>L. O. Ollila and L. A. Chamberlain</i>	

FACULTY OF EDUCATION
The University of Alberta

H. E. POOLE

Dalhousie University

A Logically *First* Stage in Training Creative Teachers?

One of the skills demanded of the creative elementary school teacher involves facility in conceiving apt and fruitful ideas for child activity. Using 60 female education students randomly assigned to four groups, three methods of improving this skill were compared with each other and with a control. Training schedules were as follows: Group D—no training; Group C—individuals preparing lists of activity ideas on set themes; Group B—as for C, but with the aid of standard written suggestions; Group A—as for B, but with group discussion.

Each experimental group spent the same amount of time in training and prepared lists of ideas over a period of five weeks on 20 themes. Two final lists of ideas prepared by all Ss were then rated by independent judges according to pre-established criteria. The predicted superiority of training over non-training groups was confirmed, but the superiority of the discussion group over both other training groups was not confirmed. Other evidence, however, provides grounds for preferring the method followed in the discussion group. It is suggested that the precepts of the "Methods" lecture need to be supplemented by this more basic form of skill learning. (Howard Poole is Assistant Professor in the Department of Education at Dalhousie University.)

The Problem

A process can only truly be described as educational when it goes beyond the training of semi-mechanical competences and seeks to develop understanding. Contemporary theory and school experience favour the view that at the elementary stage of education, in particular, learning is most likely to produce understanding when children enjoy some degree of choice in regard to what and how they shall learn. Neglect of this principle was responsible for the sheer inefficiency—as well as the drudgery—of traditional education. Children for the most part achieve insight in the course of their own purposeful activity, suitably guided. On this view, the teacher's main function is that of providing a framework within which these active experiences can occur, and yet assume significance and

coherence within the general curriculum. One aspect of this task is provision by the teacher of a choice of active experiences that are both appealing and relevant.

Many educators will agree that too few teachers possess the abilities required to meet this kind of educational challenge. One of the most commonly encountered excuses for dull classroom climates is that the teacher “couldn’t think of anything (stimulating and worthwhile) for the children to do for themselves.” Can anything be done to remedy this state of affairs? Can the teacher be helped to acquire a more imaginative appreciation of the possibilities in a learning situation? The study reported below attempts an experimental assessment of three methods of attacking this problem. These methods are adaptations of training paradigms which have long been employed in teacher education programs.

The Conceptual Framework

The ability to provide fruitful and absorbing experiences in the form of pupil activities is complex: it involves both planning and executive phases, each of which would appear to incorporate an extensive range of sub-skills. The present study will confine its attention to a single aspect of this overall planning operation. This aspect is, as the title of the report suggests, the logically basic ingredient of this enormously complex skill; namely, the ability to produce *ideas* for activities—Activity Idea Production Ability (AIPA). All teachers can be presumed to possess this ability in some degree; the problem to be investigated in this study is whether AIPA can be improved by training. The term “activity” is used here to refer to relatively transient processes in the various modes of constructing, discussing, observing, illustrating, composing and so forth which a teacher may encourage children to engage in singly or in groups as a means to promoting specifiable learning goals. The skilful exercise of AIPA will consist in producing ideas for pupil activities in suitable detail and with due regard to prevailing circumstances; in other words, although these ideas will tend to embody familiar schemas, it is only by combining and modifying them to suit particular individuals, groups, or occasions that they can be made relevant. It follows that any attempt to improve AIPA will consist not merely in instilling a repertoire of Activity Ideas, but rather in promoting an ability to modify existing repertoires creatively. Whereas repertoires can be learned by reading books about teaching, it seems doubtful in the light of teacher training practice hitherto whether AIPA can be substantially improved by this means alone.

The Present Study

The aim of the experiment described in this report was to demonstrate that ability in generating activity ideas (AIPA) can be improved by training and to determine which of the more obvious training options would result in greatest improvement. The basic training employed a variant of Maltzman’s technique (1958) for producing originality in word associations. Ss were required to produce continually different activity variations on themes provided by unit topics. Although the prescribed topics were similar in general character, Ss were instructed to avoid, as far as possible, repetition of activity ideas within and between topics.

In order to determine which treatment most improved the quality of activity idea production—which was to be judged in terms of relevance, feasibility, educational value, originality, and comprehensiveness—three experimental groups preparing unit activity lists under different conditions of enrichment were compared on terminal tasks with each other and with a control group. Enrichment was defined in terms of the volume of theme-relevant ideas available to the learner during training and/or the opportunity to discuss these ideas.

Differential enrichment was arranged as follows. In the first (non-enrichment) group, Ss were required to prepare individual unit activity lists on specified topics without faculty or other guidance and without discussion. Members of the second (reading enrichment) group were asked to prepare lists on the same topics, but received at the commencement of each session mimeographed model activity lists on the prescribed topic. Finally, Ss in a third (discussion enrichment) group, although allotted the same tasks as the other two groups and provided with models, generated their lists by small-group discussion.

On the basis of general principles of learning and the claims made by Osborne (1957) and other writers regarding the superior productivity of groups over individuals, the following hypotheses were formulated:

1. That groups practising AIPA would achieve significantly higher scores on a terminal test requiring AIPA than would be achieved by a control group.
2. That, of the groups engaged in training, those provided with enrichment would show significant terminal superiority in AIPA over the non-enrichment group.
3. That the enrichment group which combined training with small-group discussion would score significantly higher than the other enrichment group.
4. That enjoyment of training would be significantly greater in the discussion enrichment group than in the other groups.
5. That the AIPA developed in training on one sort of curriculum material would show evidence of transfer to another partially related subject matter.

The prediction contained in Hypothesis 3 was made despite the evidence (summarized by Kelley & Thibaut, 1969; Bouchard & Hare, 1970) that groups inhibit rather than promote productivity. The investigator believed that the special characteristics of the discussion group situations in this experiment would overcome the usually cited constraints placed upon group members.

Method

Subjects

Ss were 60 female Canadian elementary education students almost entirely of Nova Scotian birth and upbringing, randomly assigned in equal numbers to a control and three experimental groups. So as to avoid as far as possible the so-called Hawthorne Effect, the control group participated as Ss in another unrelated training program running concurrently and demanding exactly the same time commitment. Outside these ex-

perimental treatments, all Ss pursued exactly similar courses of instruction which unavoidably involved some unit plan preparations, and they had all previously undergone four weeks of practice teaching. Ss were told that the aim of the training sessions was to improve their lesson planning ability but were not apprised of the experimental nature of their involvement.

Procedures

Training for the three experimental groups consisted in the preparation of 20 standard Social Studies Unit Activity Lists. The emphasis to be given each unit topic was fairly exactly specified. Ss were required to list activities relevant to each topic and were continually reminded to avoid repetition of previous ideas, unless a previous idea was especially appropriate to a topic; in other words, Ss were required to justify to themselves the inclusion of previously conceived material. They were moreover instructed to concentrate on the quality and quantity of activities rather than attempt to produce coherent unit plans. Precise conventions were established for the preparation and layout of lists, and for selection of activities.¹ The topics were taken in equal numbers from the grade three and grade five Social Studies syllabuses for the Province of Nova Scotia and were administered alternately. Two lists each of 50 minutes duration were prepared during each session of one hour, with each group undergoing two sessions per week for five weeks.

The three experimental treatments were as follows:

1. Group C (non-enrichment group): individual written preparation of unit activity lists during training sessions.
2. Group B (reading enrichment group): as for 1, but Ss were provided with mimeographed models for each list they prepared. These models consisted of (a) lists of activities relevant to the topic prepared beforehand by persons not involved as subjects in the experiment and (b) the lists prepared by the discussion enrichment subgroups on the same topic.
3. Group A (discussion group): lists were prepared by three constituent groups of five persons each randomly assigned and constant in membership throughout the experiment. Each group was provided with models as mentioned in 2a) above. Ss in each group took turns in acting as recorder. Mature students uninvolved as Ss in the experiment acted as timekeepers and distributors of materials. Otherwise they detached themselves from the proceedings and offered no comments or evaluation.

In order to control as far as possible the effects due to discussion, interpersonal communication and exchange of ideas was permitted only in Group A, though no communication between subgroups of A was permitted. Furthermore, all materials and lists used during training sessions were confined to the training rooms. Lists were kept, as they were written, in individual files to which the students had free access only during training sessions. It was not possible, of course, to control out-of-session discussion among Ss or reading by those who might have been motivated to improve their performance in training sessions.

Evaluation

Quality of Activity Idea Production

1. The terminal tasks. During the week following the termination of training all experimental and control Ss were required to prepare, individually and without assistance of any kind, lists of activities relevant to two topics. The first of these was a grade four social studies topic, the second, a science topic for the same grade. No special scientific knowledge was required for this latter topic; it was employed as a terminal test in order to ascertain whether the anticipated improvement in AIPA would transfer to a different though related subject area. Forty minutes was permitted for the preparation of each list. Ss were asked to provide capsule descriptions of ten different activities relevant to the prescribed topic. All Ss made the required ten entries on both tests. Previously established conventions for plan preparation were observed. No indication had previously been given of terminal evaluation.

The resulting 1,200 activity ideas (60 Ss x 2 topics x 10 activity ideas) were corrected for grammatical and spelling errors, typed onto cards, coded for subject and treatment and then presented to three independent raters who were specialists in elementary education and strong advocates of activity methods. The social studies and science decks were evaluated separately.

2. Rating procedure and criteria. Raters were thoroughly briefed as to what was expected of them and "dry runs" indicated an initial high concordance in their ratings. The criterial attributes of "goodness" in an activity idea were:

- a. First and foremost, that it should show evidence of originality, in the sense of capitalizing upon the less obvious possibilities for pupil/group/class activity;
- b. In addition to the above, it must also be (1) relevant to the specified unit topic, (2) valuable as a learning experience for children at the intended level, (3) feasible in terms of human and material resources. The resources that Ss could take for granted in this respect were defined in "Mimeographed Instructions" (see previous footnote); accordingly they were able to assume an unusually richly provided and sympathetic school and community environment, and (4) comprehensive rather than merely a fragment of an idea; that the idea should have been developed, though in a manner not necessarily related to the length of the capsule description. In fact, a large proportion of the ideas consisted of a single short sentence.

The above criteria clearly overlap and raters therefore could only be asked to bear them in mind. For each of the two terminal decks, raters were asked to place cards bearing the activity ideas into piles according to an A to E rating (A: outstandingly original and in general keeping with the other criteria; C: not original but serviceable and in general keeping . . .; E: possessing no merit; with B and D occupying mid-positions on the scale.) Raters were warned about response sets and about the desirability of achieving something approaching a normal distribution. Although the A and E ratings were rarely used, evenness was sufficiently achieved in the other categories to permit the subsequent analysis. After the first sort, raters were asked to go through the piles again, making adjustments across rating grades. Then they were asked to go through the piles a third time for final corrections.

The aggregate score (possible 40, A=4, E=0) awarded by each rater to each activity list (ressuscitated by drawing together ten component scores) was computed and combined with the other two ratings (interpreted as scores) for the same list. A final score for each subject (out of 120) was thus arrived at for each of two terminal tests.

Results

The validity of conclusions drawn in this report depends heavily upon there being an acceptable degree of concordance between raters. Tables 1A and 1B provide a measure of reliability of raters on both terminal tasks in terms of intraclass correlation.

TABLE 1A
RELIABILITY OF RATERS BY INTRACLASS CORRELATION—
SOCIAL STUDIES LISTS

Source of Variation	SS	df	MS	F	
Persons	3940	59	66.7	7.7	$p < .001$
Raters	8	2	4.0	1	N.S.
Remainder	969	118	8.2		
Total	4917	179			

Reliability for a single rater: 70.
Reliability of mean of 3 ratings for each person: 87.

TABLE 1B
RELIABILITY OF RATERS BY INTRACLASS CORRELATION
SCIENCE LISTS

Source of Variation	SS	df	MS	F	
Persons	4395	59	74.5	8.09	$p < .001$
Raters	14	2	7.00	1	N.S.
Remainder	1088	118	9.2		
Total	5497	179			

Reliability for a single rater: 70.
Reliability of mean of 3 ratings for each person: 88.

The absence of significant difference between raters in spite of the highly significant differences between Ss, provides justification for summing the ratings (interpreted as scores—the rationale for using parametric tests on these data is argued by Anderson, 1961) awarded by each rater to each S in order to yield the data in Table 2. The control group is designated D.

TABLE 2
MEAN SCORES ON TWO TERMINAL TESTS BASED ON SUMMED RATINGS
FOR EACH ITEM (POSSIBLE 120)

Group/Treatment	(1) Social Studies		(2) Science	
	Mean	SD	Mean	SD
A	80.1	14.45	75.8	13.2
B	74.4	12.7	72.7	12.4
C	68.7	10.6	67.0	10.5
D	65.6	12.8	64.6	11.1

This array of data permits the test of Hypothesis 5 (the transfer hypothesis). The science activities task required Ss to exercise AIPA in a subject area different from that involved in training. It is assumed that some adaptation of previous learning was thereby demanded, although logical similarities between the structure of science and social studies at the grade level in question ensure that many activity ideas are mutually applicable.

The hypothesis is tested in the following manner. If transfer occurs as predicted, the *rank order* of mean scores on the second (science) task will resemble the rank order of the same groups on the first task. A perfect correspondence in rank order proves to be the case. But the probability of correctly predicting this permutation is $1/4!$ Accordingly, the transfer hypothesis is confirmed beyond the 95% confidence limit ($p < 0.042$). Whether this effect would carry over to areas of widely differing epistemic structure (e.g., mathematics) is of course unsure on this evidence, though some Ss have reported such overall improvement in AIPA as a result of training.

Next, the data from which Table 2 was compiled are tested below for significant differences between groups.

TABLE 3
EFFECTS OF PRACTICE
SUMMARY OF ANALYSIS OF VARIANCE OF SCORES ON TWO
TERMINAL TESTS

Source of Variation	(1) Social Studies				(2) Science			
	SS	df	MS	F	SS	df	MS	F
Between Groups	1841	3	613	3.5**	1183	3	394.3	2.6*
Within Groups	9819	56	715.3		8487	56	151.5	
Total	11659				9670			

** $p < .025$

* $p < .10$

A significant result in 3(1) justifies a preplanned comparison of treatments.

TABLE 4
COMPARISON OF TREATMENT EFFECTS UPON SOCIAL STUDIES SCORES:
PARTITIONING OF SS_{bg} FROM TABLE 3(1)

Source of Variation	SS	df	MS	F	p
Between Control & Treatment Groups (D vs. A, B, and C)	866	1	866	4.93	< .05
Between Enrichment & Non-Enrichment (C vs. A and B)	735	1	735	4.19	< .05
Between Kinds of Enrichment (A vs. B)	240	1	240	1.37	N.S.

Hypothesis 1 (effect of practice) and Hypothesis 2 (effect of enriched practice) are both confirmed, but the test of Hypothesis 3 (superiority of discussion over reading enrichment) does not achieve significance.

The prediction of greater enjoyment by the discussion treatment (Hypothesis 4) is tested by organizing the confidential evaluations of training made by Ss in Table 5. The evaluations took place immediately after the terminal tests.

TABLE 5
ENJOYMENT OF TRAINING BY TREATMENT GROUP¹

Group	Response Choices			
	High Enjoyment	Low Enjoyment		
	Very Much	Moderately	Very Little	Not at All
Discussion A	7	3	3	2
Non-Discussion B	0	6	6	3
C	1	4	5	5

¹Frequency of qualification chosen to complete statement beginning (I enjoyed the training sessions..."

The hypothesis is tested by pooling the frequencies in the non-discussion and low-enjoyment categories to produce a 2 x 2 table of which $\chi^2 = 14.6$ with 1 d.f.: $p < 0.001$ (one-tailed test).

Finally, the data presented in Table 6 are relevant to establishing the claim that *originality*, as well as recollected activity ideas, has contributed to the scores of treatment groups. Resolution of this point is crucial to the basic position taken in this report—that the higher scores of experimental Ss is evidence of a more original approach to unit planning. The problem arises in this way. Activity ideas may be original productions or they may be items recollected from previous experience or training sessions. Clearly there can be no certain way of deciding to which category a particular idea belongs. In this sort of situation it is tempting to apply Occam's razor—that is to say, if the terminal performances are sufficiently explained by recollection of previously en-

countered ideas, it is both unnecessary and illegitimate to postulate that *originality* as well as *repertoire* has been learned.

The problem is aggravated by the difficulty of providing an external criterion for the existence of the originality component in AIPA. This would entail showing that the training described in this report would significantly affect scores in tests of originality or creativity which were general in character. With this aim in view the present author, having previously trained 20 Ss on Treatment A—though with much more involvement in the proceedings on his part—and having discovered significant differences in quality of activity idea production between this group and a control, administered the Torrance Tests of Creative Thinking to both groups. No significant difference between groups on the TTCT was evident. The investigator concluded that no simple identification of AIPA with creativity was warranted.

It was decided therefore to use a weaker, internal criterion of the predicted increase in originality. Improved originality would be assessed relative to the body of responses actually generated by the four groups on the terminal tests. This procedure certainly restricts the extent to which the results of this study can be generalized, but it seemed the only available strategy in the circumstances. Accordingly, the frequency of A ratings awarded by the raters on each test were extracted from the data and summed to produce Table 6. These “outstandingly original” ratings were few in number—less than two percent of the total number of ratings—which, for reasons mentioned by Guilford (1954, p. 297) in connection with ratings of humour, is hardly surprising. After three sortings, initial judgments of originality tend to be revised downwards.

TABLE 6
ORIGINALITY IN ACTIVITY IDEA PRODUCTION¹

	Group			
	A	B	C	D
Summed Frequency of 'A' Ratings on Two Terminal Tests	26	19	11	13

$$\chi^2 = 7.935 \text{ with } 3 \text{ df} \quad p < .05 \text{ (One-tailed test)}$$

¹Goodness of fit of frequency of 'A' ratings by group to a rectangular distribution.

If originality had not been learned during training, the A ratings should have been evenly distributed among the four groups. Table 6 shows that this is not the case. Despite a slight variation in the readiness of raters to award A's, the results are indicative of a significant trend in the predicted direction.

Discussion

Of the five hypotheses tested, four have been confirmed at high or acceptable levels of confidence. Comment will be restricted to the failure to confirm Hypothesis 3. The superiority of discussion over reading enrichment although in the predicted direction, was not borne out at the predetermined .05 level of significance. The author is strongly inclined to attribute this failure to the very attempt to control conditions in Treatment A which was a prerequisite for valid subsequent comparison of experimental treatments. On the basis of his per-

sonal experience with groups, the author had believed that the advantages of groups over individuals, as detailed by Maier (1967), would be maximized when groups consisting of teachers, or intending teachers previously exposed to classroom situations, were confronted with the task of co-operatively planning classroom tactics. It had always been the author's impression that elementary school teachers were unusually ready, as compared with other occupational groups, to instruct others, including associates, on matters pertaining to teaching. Furthermore, it was thought that groups would be more apt than individuals to avoid repetition of previous ideas and thus provide a greater challenge to innovation.

However, the level of interaction in the three subgroups of Treatment A departed considerably from expectations. One group was extremely productive, another enjoyed spasmodic periods of inspiration, while the remaining group performed rather sluggishly throughout. The consequent variation in the training environment for members of the subgroups accounts for the relatively large variance in A group scores on the social studies test (Table 2) and is consistent with the interpretation invited by Table 5 that the benefit derived by different individuals from Treatment A varied greatly.

In summary, the following explanation is offered for the failure of Treatment A to achieve its objectives relative to Treatment B: that the level of maturity and professional preparation of the Ss involved in the discussion group was insufficient to enable them to function as autonomous groups; that many Ss at this stage of teacher development require motivation and leadership from outside the group. Unfortunately, this latter sort of influence upon the dynamics of a group does not easily lend itself to the sort of analysis attempted in this study. The confirmation of Hypothesis 4 (that discussion will be the most enjoyed treatment—Table 5) does, however, provide a reason of a different kind for preferring discussion to the mere assignment of reading material in training AIPA.

Applications and Conclusion

The success of the child-centred movement in education depends upon higher levels of instructional skill than have previously been demanded of teachers. One important ingredient of this skill is AIPA. Being able to conceive clever, original ideas for pupil activity is a necessary—though not a sufficient—condition for successfully managing learning in a modern elementary school. Yet there are still many educational systems and cultural settings which, in one way or another, discourage teachers from developing or exercising AIPA. The consequence of this deficiency must be a less varied learning environment for children.

AIPA training is the beginning of an answer to this problem. It is an attempt to create a partial bridge between the somewhat abstract counsels of the Methods Lecture, on the one hand, and Teaching Practice on the other. Because AIPA training involves the intending teacher in *activity*, it is assumed to produce learning at a more fundamental level than would be expected from the digestion of verbal precepts alone. The present study has shown that AIPA can be improved by practice, especially by practice under enriched conditions. Although the relative merit of the two enrichment conditions has not been conclusively determined, the greater enjoyability of the discussion

A Logically First Stage in Training Creative Teachers?

method (Table 5) recommends its use over that of a regime based solely upon printed matter.

The most promising application of these findings would seem to be in the incorporation of small group techniques within the more traditional patterns of Methods courses rather than by employing the actual schedules of group discussion described in this report. In conclusion, the present writer believes that further isolation and study of the trainability of component skills of teaching could contribute towards the improvement of teacher education programs.

The author acknowledges his gratitude to Dalhousie University for the facilities and grants which made this study possible.

¹ Mimeographed "Detailed Instructions for List Preparation" are obtainable from the author. The most important effect of these instructions was to control the complexity of the context in which AIPA was exercised. This was achieved by making as many features as possible of the hypothetical school situation (age, ability of children, school characteristics, etc.) explicit and standard for all Ss.

References

- Anderson, N. H. Scales and statistics: Parametric and nonparametric. *Psychological Bulletin*, 1961, 58, 305-316.
- Bouchard, T. J., & Hare, M. Size, performance and potential in brainstorming groups. *Journal of Applied Psychology*, 1970, 54 (1), 51-55.
- Guilford, J. P. *Psychometric methods*. New York: McGraw Hill Co., 1954.
- Kelley, H. H., & Thibaut, J. W. Group problem solving. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology*, Vol. IV, (2nd ed.). Reading, Mass.: Addison Wesley Publishing Co., 1969, 1-101.
- Maier, N. R. F. Assets and liabilities in group problem solving: The need for an integrative function. *Psychological Review*, 1967, 74, 239-249.
- Maltzman, I., Bogartz, W., & Breger, L. A procedure for increasing word association originality and its transfer effects. *Journal of Experimental Psychology*, 1958, 56 (5), 392-398.
- Osborne, A. F. *Applied imagination*. New York: Scribner, 1957.

T. O. MAGUIRE

University of Alberta

R. B. PATSULA

Edmonton Separate School Board

P. O. EVANECHKO

University of Victoria

The Development of Word Meaning Discrimination in Children

In the past, investigations have been made into the development of word definition abilities in children. Such studies focused on the child's ability to formulate formal definitions. In the present study, the development of the ability to differentiate among various kinds of word meaning was explored. The results indicated that children in the eighth and eleventh grades (about 13 to 16 years of age) are capable of discriminating among various kinds of word meaning. The consequences of the research for the development of achievement tests is discussed. (Dr. Maguire is Professor of Educational Psychology at The University of Alberta; Mr. Patsula is a teacher with the Edmonton Separate Schools; Dr. Evanechko is Associate Professor in the Faculty of Education, the University of Victoria.)

Several researchers have investigated the development of word definition abilities in children (Wolman & Barker, 1965; Al-Issa, 1969; Swartz & Hall, 1972). Their results are consistent in indicating that as children grow older, their choice of definitions pass through three general stages: descriptive (an apple is red), functional (an apple is good to eat), and categorical or abstract (an apple is a fruit). While these findings are interesting relative to the development of formal operations in children, they do not cast much light on the broader question of the development of a differentiated concept of word meaning. In short, previous research has been concerned with the question, "When do children begin to develop formal definitions?"; the broader question is, "When do children begin to realize that there are different kinds of meaning?"

Despite its age, the mediational view of meaning outlined by Osgood, Suci, and Tannenbaun (1957) provides a useful model of meaning with the idea that a word derives its meaning by being embedded in an individual's semantic space. Semantic space is a term used to describe the network of relationships

which link internal mediating responses. The position of a word in the semantic space in effect describes its meaning. In their studies of the connotative meaning of words, Osgood and his associates produce a dimensional analysis of connotative links that exist in the semantic space. There are many other kinds of links that exist, however, and in the present study, attention was directed toward the emergence of these links in children.

Evanechko and Maguire (1972) suggested that the semantic space is comprised of 24 kinds of logico-semantic relations which in effect are ways in which words possess meaning. They attempted to resolve the 24 kinds of meaning into a smaller set of dimensions to see if the resulting configurations, when derived for children at two age levels, indicated a developmental trend. They suggested that differences between children in grades 5 and 8 existed in the semantic structures, with younger children having spaces oriented toward experience, and older children showing more sophisticated class structures. These results were in agreement with the earlier research on word definitions, but again, no attempt was made to address the broader question of whether the older children possessed more differentiated semantic structures.

According to the mediational view, as the child grows older and gains more experience with his verbal environment, there should be a concomitant increase in the number of ways in which he can attach meaning to words. The purpose of the present study was to investigate this development by examining the ability to discriminate among kinds of word meaning in children at three age levels.

The logico-semantic relations used by Evanechko and Maguire (1972) were derived by Evanechko (1970) from the literature on the development of children's word definitions (Annett, 1959; Burns, 1960; Cronbach, 1943; Dale, Eicholz, & Bennet, 1960; Fiefel & Lorge, 1950; Flavell & Flavell, 1959; Lewinski, 1948; Petty, Herold, & Stoll, 1968; Russell, 1954; Vinacke, 1951; Welch, 1940). The 24 links, together with an example of each, are shown in Table 1. They should be taken as neither strictly mutually exclusive nor exhaustive.

In the present study, the children were asked to sort several examples of different meaning types into categories, and then to explain the strategy that was used for any one of the categories. The work of Vygotsky (1962) and Bruner (1964) provided the framework for interpreting the sorting strategies.

Vygotsky (1962) takes the position that three phases are passed through in the ascent to concept formation.

1. Objects are placed together in heaps for purely subjective reasons.
2. Objects are placed together according to objective bonds that exist between them, but these bonds lack a logical unity so that there may be many different rules for including the different objects in a single group. The objects may even be placed together by connecting them with a story or theme. The groups at this stage are called complexes.
3. Objects are placed into groups according to a single rule.

Bruner and Oliver (1963) and Bruner (1964) noted a similar developmental trend in that younger children rely most heavily on perceptual attributes of objects to form complexes, but as they grow older they begin to focus on functional properties to form superordinate groupings, or true concepts. Both the Bruner and Vygotsky views are consistent with the research on the development of word definition abilities cited earlier. In the present study, an

attempt was made to classify the sorting strategies of the subjects in accordance with the developmental sequence outlined above.

Method

The Task

A pilot study was carried out using three decks of 48 cards, each card illustrating one relationship (e.g. rob-steal). Each deck contained two examples from each of Evanechko's 24 categories. The subjects were asked to verbalize their grouping strategies. From the results, it was found that the task was too unwieldy, particularly for younger subjects. The subjects required a huge physical area to do the task; they could not keep track of their reasons for sorting; and with only two examples for each category they found it difficult to obtain satisfactory closure on their categories.

It was decided to alleviate the problem by increasing the number of examples available from each theoretical category, and decreasing the number of categories represented within the set that a particular individual had to sort. Ultimately it was decided to break the 24 categories into five sets (one set of four categories and the rest having five categories). The disadvantage of this technique is that it is impossible to see if the children can discriminate among all of the categories at the same time. Because of this, it was decided to try to place the categories into sets in such a way that categories which might easily be confused with each other should appear in the same set, and that categories with obvious differences should be separated. In other words, steps had to be taken to maximize the opportunities of rejecting the hypothesized structure.

Since Evanechko (1970) had suggested five logical sets of the categories, this provided a useful starting point. Pilot studies revealed that in simplifying the sorting task most individuals divided the examples into two piles, one containing categories with single word definers, and the other containing categories of multiple word definers. Since the subjects did this anyway, care was taken to place single word and multiple word definers in different sets. Also, in the Evanechko and Maguire study as well as in the pilot work it was found that subjects confused Superordinate and Generic meanings, Coordinate, Part-Part and Free Association meanings, Attribute and Action-of meanings, and Repetition and Denotation in Context meanings. In grouping the categories, often confused categories were placed in the same set. The resulting sets are shown in Table 1.

TABLE 1
TWENTY-FOUR CATEGORIES OF LOGICO-SEMANTIC RELATIONS
EXHIBITED IN DEFINITIONS USED BY CHILDREN

Set	Category
A	1. <i>Synonym</i> . The members of each word pair have exactly or very nearly the same referent: e.g. steal—rob
A	2. <i>Similarity</i> . The members of each word pair are similar through being aligned on some dimension, with the referent of the right-hand member occupying a more extreme position of this dimension: e.g. hungry—starving

The Development of Word Meaning Discrimination in Children

- A 3. *Superordinate*. The left-hand member denotes a common class of which the right-hand concept is a member:
e.g. fruit—apple
- A 4. *Whole-part*. The right-hand member of each pair refers to a familiar object recognized as an important part of a familiar whole denoted by the left-hand member:
e.g. bird—wing
- A 5. *Generic Definitions*. The right-hand member denotes the common class to which the left-hand member belongs:
e.g. cup—dinnerware
- B 6. *Coordinate*. The members of each pair refer to familiar members of a familiar class:
e.g. chair—table
- B 7. *Contrast*. The members of each word pair refer to opposite ends of a continuum:
e.g. hard—easy
- B 8. *Part-part*. The members of each pair refer to familiar objects which are parts of a familiar whole:
e.g. wall—floor
- B 9. *Free Association*. The members of the unit are free associates:
e.g. carry—heavy
- B 10. *Connotation*. The right-hand member of each pair connotes a relationship with the left-hand member:
e.g. royal—strong
- C 11. *Attribute*. The right-hand member of each pair refers to a quality or attribute generally recognized as characterizing the object denoted by the left-hand member:
e.g. turtle—slow
- C 12. *Action-of*. The right-hand member of each pair is an intransitive verb denoting concrete action associated with and performed by the agent referred to by the left-hand member:
e.g. baby—cry
- C 13. *Action-upon*. The left-hand member of each pair is a transitive verb denoting a concrete action associated with and performed upon the object referred to by the right-hand member:
e.g. sweep—floor
- C 14. *Common Use*. The right-hand member of each pair denotes an object associated with and acted upon by the agent referred to by the left-hand member:
e.g. farmer—tractor
- D 15. *Use Of*. The right-hand member of each unit denotes a use made of the left-hand member:
e.g. envelope—for putting letters in
- D 16. *Contiguity*. The left-hand member of the unit is defined by direct concrete interaction of place, time or activity with the right-hand member.
e.g. late—you can see by the clock
- D 17. *Analysis*. The right-hand member is an analysis of the left-hand member indicating certain dimensions of function of this concept:
e.g. lengthen—make a thing longer
- D 18. *Synthesis*. The right-hand member defines the left-hand member by stating its relation with other concepts commonly associated with it:
e.g. bunk—it has two levels
- D 19. *Ostensive Definition*. The right-hand member defines the left-hand member largely on the basis of experience:
e.g. selfish—all for yourself

- E 20. *Repetition*. The right-hand member of each unit is a repetition of the concept referred to by the left-hand member:
e.g. drink—a drink of water
- E 21. *Extension of a Class (Implication)*. The right-hand member of the unit gives examples of concepts to which the left-hand member might refer implying a degree of familiarity with the concept:
e.g. farming—crops and animals
- E 22. *Denotation in Context*. The left-hand member is defined by use in context:
e.g. sharpen—sharpen the knife till it cuts well
- E 23. *Class Membership Implied*. The right-hand phrase attempts to bridge the gap between general and specific by using phrases such as “a kind of”, “sort of” or “like a”:
e.g. stool—like a chair
- E 24. *Intension of a Class (Genus et Differentia)*. The right-hand member states the class as well as the distinguishing features of the left-hand member:
e.g. sipped—drank a little at a time

For each category in each set, six examples were selected from Evanechko's pool, so that each set of categories consisted of 30 examples (except for C which had 24). The definitions were printed on two inch by four inch cards and randomly arranged in decks corresponding to the five sets.

Subjects

The task was administered to 176 subjects in grade 5, 196 subjects in grade 8, and 198 subjects in grade 11. An approximately equal number of subjects sorted each set of items at each grade level. Each subject was given a copy of the introduction, an answer sheet, and one of the five sets of cards. Care was taken to insure that neighbouring students received different decks of cards. In each set, the cards were arranged in random order.

Procedure

An introduction was given consisting of two parts. In the first part, classification principles were illustrated using geometric figures of different sizes. The students were shown that there are many ways of grouping the figures and that any way is correct provided that there is some reason for putting an object in a group. In the second part of the introduction, a set of nine definitions not used in the study was grouped for the students in two different ways. Care was taken to indicate that the illustrated groupings were not the only ways that the definitions could be grouped. Although there was some concern that the example sorts might influence the grouping strategies of the subjects, it was decided that if the Vygotsky and Bruner positions had much validity, the strategies employed by the subjects would not be greatly influenced by the two examples. In short, the subjects would hear what they were capable of hearing. The advantage of the two examples was to show the students the form of the response required, and to illustrate that there was more than one way to do the task. The students were told that they could use as many categories as they wanted to, and that they could put as many cards in each category as they wanted to, but that all cards in a set should have similar definitions.

Method of Analysis

The data were analyzed for each grade by set combination by the latent partition analysis (LPA) procedure outlined by Wiley (1967). In general, LPA is used in situations in which subjects partition a set of items into a number of

categories, where there are no restrictions on the number of partitions, nor on the number of items placed in each of the categories. The LPA model assumes that when a relatively homogeneous group of people sorts a set of objects, there exists a latent partitioning of the objects which underlies each individual's manifest partition. In the hypothetical case different manifest partitions arise from the combination of various latent categories, or from the fractionation of latent categories. The basic model is

$$S = \phi' \Omega \phi + \Delta^2.$$

Where S is a matrix of item joint occurrences (the proportion of times pairs of items are sorted together), ϕ is the latent partition matrix which in the errorless case consists of 1's and 0's according to whether an item is in a particular latent category or not, Ω is the confusion matrix indicating the probability of an item being included in two different categories under independent sortings, and Δ^2 , is the probability of items being included in two different categories under independent sortings.

The procedures for determining the number of latent categories and the matrices ϕ , Ω , and Δ^2 , are described in Wiley (1967), and need not be reiterated. In the present study, for each set of definitions, there are four partitions available, ϕ_{11} , ϕ_8 , ϕ_5 , ϕ_T . The subscripts 11, 8, and 5 refer to grade level; the subscript T refers to the theoretical partition based on Evanechko's categories of definitions. The ϕ matrices for grade 5, 8, and 11 were cross tabulated with each other and with ϕ_T for each set.

The extent to which the four partitions of each of the five sets agreed with each other was measured using Evan's (1970) index of agreement A . This index ranges from 0, representing no agreement to 1 representing complete agreement. If X_{ab} is a matrix whose elements x_{jk} are the numbers of times examples of definitions in the j^{th} category of partition 'a' also occur in the k^{th} category of partition 'b', then a perfect cross tabulation exists between a and b if X_{ab} is a square matrix with only one entry in any row or column being greater than zero. For these partitions $A = 1$. For less perfect agreement the calculation of A depends on the amount of disagreement relative to the maximum amount of disagreement. The calculation of the maximum amount of disagreement described by Evans was modified by Patsula (1972) to avoid situations in which the maximum possible disagreement as defined by Evans is based on situations that can never occur.

Results of the Latent Partition Analysis

The indices of agreement between the latent partitions for each of the five sets are shown in Table 2. In Table 3, the cross tabulation matrices between each grade and the theoretical partition are presented along with the index of agreement and the number of subjects upon whom the empirical partitions are based. From these tables it can be seen that in all sets except C the degree of agreement between the latent partitions and the theoretical partition increased from grade 5 to grade 8 to grade 11. In Set C, the agreement between T and 11 is less than for either 8 or 5. If one postulates a developmental trend toward T, then the data are supportive of the hypothesis in all cases except C. Since the sampling distribution of the index of agreement is not known, it is impossible to investigate the differences on the basis of inferential statistics.

Considering Table 3, it must be admitted that the differences in contingency tables whose indices of agreement are of the order of .72 to .76 (Set B, grades 5 and 8) would be difficult to detect by eye. Differences of the order .79 to .91 (Set C, grades 11 and 8) are more obvious.

TABLE 2
INDEX OF AGREEMENTS BETWEEN DIFFERENT PARTITIONS
OF THE FIVE SETS OF DEFINITIONS

Between	A	B	Set C	D	E
T & 11	0.94	0.84	0.79	0.94	0.80
T & 8	0.74	0.76	0.91	0.75	0.61
T & 5	0.62	0.72	0.88	0.55	0.46
11 & 8	0.79	0.71	0.90	0.79	0.69
8 & 5	0.72	0.74	0.84	0.76	0.62
11 & 5	0.59	0.72	0.74	0.57	0.56

TABLE 3
CROSS TABULATIONS BETWEEN THE LATENT PARTITIONS AT EACH
GRADE LEVEL AND THE THEORETICAL PATTERN

Grade	Set A	Set B	Theoretical Set C	Set D	Set E
11	6 0 0 0 0 0 6 0 0 0 0 0 6 0 0 0 0 0 6 0 0 0 0 0 3 0 0 0 0 3 A = .94 n = 40	6 0 0 0 0 0 6 0 0 0 0 0 6 4 0 0 0 0 2 0 0 0 0 0 4 0 0 0 0 2 A = .84 n = 39	6 0 0 0 0 6 0 0 0 0 4 0 0 0 0 3 0 0 2 3 A = .79 n = 40	6 0 0 0 0 0 6 0 0 0 0 0 6 0 0 0 0 0 6 0 0 0 0 0 3 0 0 0 0 3 A = .94 n = 39	6 1 0 0 0 0 5 0 0 0 0 0 4 0 1 0 0 2 1 1 0 0 0 5 0 0 0 0 0 4 A = .80 n = 40
8	6 0 0 0 0 0 5 1 2 0 0 0 5 0 0 0 1 0 3 0 0 0 0 0 3 0 0 0 1 3 A = .74 n = 40	6 0 0 0 0 0 5 0 0 4 0 0 4 0 0 0 0 2 1 0 0 0 0 5 0 0 1 0 0 2 A = .76 n = 38	6 0 0 0 0 6 0 0 0 0 6 0 0 0 0 3 0 0 0 3 A = .91 n = 41	6 0 0 0 0 0 5 0 0 1 0 0 2 0 0 0 0 4 2 0 0 0 0 4 0 0 1 0 4 0 0 0 0 0 2 A = .74 n = 39	5 5 0 0 0 0 0 5 0 2 0 0 0 4 0 0 0 0 0 3 0 0 1 1 0 1 1 0 1 1 A = .61 n = 38
5	6 0 0 0 0 0 4 0 1 0 0 0 3 0 1 0 0 1 0 0 0 0 1 2 0 0 1 0 2 0 0 0 1 0 4 0 1 0 1 1 A = .62 n = 35	3 0 0 0 0 3 0 0 0 0 0 5 0 0 0 0 0 6 1 0 0 0 0 4 0 0 0 0 0 2 0 0 0 0 2 0 1 0 1 2 A = .72 n = 34	6 0 0 0 0 4 0 0 0 2 0 1 0 0 6 0 0 0 0 5 A = .88 n = 35	5 0 0 0 0 0 5 0 0 3 0 0 1 0 0 0 0 2 1 0 0 0 2 2 0 0 0 1 3 0 0 1 0 0 2 1 0 0 0 1 A = .55 n = 36	2 1 0 1 0 2 4 0 0 0 0 0 3 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 1 0 1 0 1 2 0 0 0 0 0 4 1 1 1 0 2 A = .46 n = 36

Although the differences in successive indices are not always large, the consistency of the trend (excepting Set C) which exists in independent samples with independent stimuli gives strong support to the developmental hypothesis. Additional evidence (although not independent) can be garnered from the indices of agreement between grades within sets. If a trend exists, the agreement between grades 5 and 11 should be less than either the agreement between grades 5 and 8, or 8 and 11. In all sets except B this is the case. In Set B, $A_{8,11} = .71$ and $A_{5,11} = .72$.

An interesting result that occurred was that apart from Set B, $A_{5,8}$ is less than $A_{8,11}$. This would be consistent with the idea that greater change takes place between grades 5 and 8 than from grades 8 and 11.

Both of the anomalous sets (B and C) produced fairly high values for the index of agreement (all above .7) at all grade levels. This suggests that the subjects may have found the stimuli so easy to categorize that there was no discrimination along the developmental continuum. A close examination of the groupings made in Set C by the grade 11 students indicated that their lower agreement with T was caused by a tendency to overclassify. In the Action-of category, they discriminated between auditory actions and physical actions (example: baby-cry vs. rabbit-hop). It also appears that grade 11 sorters discriminated between Attributes of inanimate and animate objects (example: flame-hot vs. turtle-slow).

The overall consistency among subjects in grouping elements can be described by the average of $(1-\delta_j^2)$ where δ_j^2 is the diversity of item j , the probability of item j being included in two different manifest categories under independent sortings. Table 4 shows the average value of δ_j^2 for the fifteen situations (five sets by three grades). Also shown in the table are the number of

TABLE 4
AVERAGE OF $(1-\delta_j^2)$ FOR THE FIFTEEN LATENT PARTITIONS

Set	Grade	Average of $1 - \delta_j^2$	Number of Theoretical Categories	Number of Categories in Latent Partition	Average Number of Categories in Manifest Partition	Standard Deviation
A	5	.605	5	8	7.5	2.68
	8	.681	5	6	5.7	2.41
	11	.727	5	6	5.6	2.21
B	5	.552	5	8	7.1	3.11
	8	.635	5	6	5.8	2.57
	11	.731	5	6	5.4	2.05
C	5	.647	4	5	6.0	2.32
	8	.822	4	5	4.6	1.14
	11	.819	4	5	5.0	1.34
D	5	.500	5	8	8.4	4.47
	8	.652	5	7	6.6	2.55
	11	.665	5	6	6.2	2.10
E	5	.553	5	9	8.1	4.47
	8	.564	5	6	6.1	3.10
	11	.623	5	6	6.2	3.22

categories in the derived latent partition, and the average number of categories used by the sorters as they did the task and the standard deviation of the number of manifest categories used.

In all cases except Set C the consistency increases from grades 5 to 8 to 11. In Set C the average value for $(1-\delta_i^2)$ is .822 for grade 8, and .819 for grade 11. This difference is very small, and both values indicate that the consistency is so high that both groups can only be described as very consistent. In all cases the average number of manifest categories is greater in grade 5 than in either grades 8 or 11. Consistent with this (and partly as a result of it), the standard deviations are also larger for the grade 5 groups, indicating less consistency in the number of manifest categories used at grade 5. When comparing the grade 8 and 11 groups, the data on the number of manifest categories does not exhibit the same strong trend. The differences between the two grades are generally small, indicating that the major differences occurred between grades 5 and 8.

Identification of Strategies

In order to investigate grouping strategies, the students were asked to pick one of their groups, and explain how the definitions were similar. As stated earlier, heaps, complexes and superordinate groupings were seen by Vygotsky and Bruner as typical stages in the development of concept formation abilities. In Table 5 the responses are grouped according to these stages, and according to specific strategies within stages.

TABLE 5
CLASSIFICATION OF GROUPING STRATEGIES

Grade	Left Blank or Meaningless Heaps	Complexes				Superordinate Concepts		Total
		Thematic Groupings	Associative Group	Key Rings	Multiple Groupings	Subject Oriented	Definition Oriented	
5	17	5	50	2	4	37	61	176
8	7	0	29	1	0	29	130	196
11	9	1	9	0	0	38	141	198

Heaps

Only five subjects, all from grade 5 did not respond to the task of selecting and describing a group of definitions. In addition, 28 subjects' answers were judged to be "heaps" or impossible to categorize. (e.g. "I put these together because they all go together.")

Grouping in Complexes

An initial phase on the way to concept formation comprises many variations of grouping in complexes. The simplest complex is *thematic grouping* in which definitions or key words are put together by virtue of participating in a sentence or a little story. In thematic grouping, n different rules are used to

account for the n different stimuli placed together in the group. An example of grouping by theme taken from the grade 5 responses to Set E is:

crime—stealing or killing
surrender—surrender or be caught and killed
sharpen—sharpen the knife till it cuts well
bloody—a bloody knife

The explanation provided by the student was:

One man is stealing and he is killing and the police said surrender because he was killing people with sharp knives and it was bloody.

Three types of grouping by complexes that are more advanced than thematic grouping are associative groups, key ring groups and multiple groups. In an *associative group*, the child notes a factual relationship between two definitions and all other definitions are related to this relationship. An example of an associative group from grade 8, Set C was:

diamonds—expensive
farmer—tractor
pilot—airplane
play—piano

The explanation given was:

Because they are all very expensive and they cost a lot to own and operate.

In *key ring* strategies, one definition is taken as the key and all other definitions in the group are included because they possess an attribute in common with the key definition. An example of a key ring strategy from grade 5, Set D was:

interfere—when you get in the way (KEY)
pebble—found lying on the ground
iceberg—from a glacier

The explanation given was:

I picked set 7 the first card I got was interfere I also picked pebble because I thought a pebble would interfere, I picked iceberg because icebergs interfere with ships because ships sometimes crash into icebergs.

In *multiple groupings*, several rules were used to justify the inclusion of different definitions. Often, the child appeared to have clustered several groups for no apparent reason. Within each of the subgroups, there was some commonality. None of the explanations were particularly helpful in providing insight into the child's strategy. While multiple groupings have been ranked here alongside associative groups and key rings, they may represent a kind of transition between the use of heaps and the use of complexes. An example of multiple groupings that appeared for grade 5, Set D was:

egg—from a chicken
stove—found in a kitchen
knife—it has a blade
scissors—for cutting
envelope—for putting letters in
fishhook—it is attached to a line and rod
apple—grown on a tree

The explanation was:

Scissors are very sharp, a knife can cut you, a stove is very hot. Envelopes are to put letters in, fishhooks are to put on lines. Apples grow, chickens lay eggs.

Superordinate Concepts

Considerably more sophisticated than complexes are superordinate concepts in which one universal rule for inclusion accounts for all of the definitions in the set. Both Vygotsky (1962) and Bruner and Oliver (1963), indicate that true concepts or superordinate concepts could be formed at various levels of sophistication. In the present study, two levels were noted. At the lower level, the subjects focussed on the perceptual or subject features of the cards. At the higher level, attention was direct to the functional or definition properties.

In *subject oriented concepts*, the word being defined or a key word from the body of the definition is taken and similarity is judged by focusing only on the quality or attributes of this single word. The sorter concentrates on the type of word being defined and placed the cards into sets of similar subject matter. An example of subject oriented grouping taken from grade 11, Set D was:

pebble—found lying on the ground
iceberg—from a glacier
dock—where whip ties up
fishhook—it is attached to a line and rod

The explanation was:

In set 4 the rule used was that the things listed had to do with land and or water.

In *definition oriented* concepts the sorter concentrates on the whole definition and places the cards into sets so that all cards in a set have definitions that are similar. The universal rule of inclusion accounts for all definitions in the set by relating how the defining words give meaning to the respective subjects of the definitions in similar manner. An example of definition oriented grouping taken from grade 5, Set A was:

dog—collie
basketball—game
apple—fruit
beetle—insect
daffodil—flower
mushroom—plant
dessert—pie
fish—salmon
bird—robin
vegetable—carrot
animal—deer

The explanation was:

In one of my sets (set 2) which consisted of 11 cards I used this rule: I looked at both words and if one of the words was a type of the other (eg. bird—budgie) I would pick it for this set.

The results shown in Table 5 indicate that the proportion of subjects using definition oriented concepts was greatest at grade 11 and lowest at grade 5. The proportion of subjects using complexes was greatest at grade 5 and lowest at

grade 11. The number of subjects using pre-concept heaps is fairly equal at all grade levels (12, 7, and 9 for grades 5, 8, and 11 respectively), but when the subjects who did not carry out this part of the task are included with this group, the number in grade five increases. Regardless of the validity of putting the non-respondents with the "heaps", the data for complexes and concepts show a strong order in favor of the hypothesis that older subjects use more sophisticated concepts in their grouping strategies.

Discussion

The results of the present study suggest that sometime between grades 8 and 11 (about 13 to 16 years of age), the students as a group become capable of distinguishing among the various types of word meaning. Although it was not possible to have subjects sort examples of all 24 relationships at one time, an attempt was made to have groups of subjects work with relationships that were likely to be confused with each other. Within this constraint, the older subjects produced latent categories that were remarkably similar to the "theoretical" categories. In addition, there was an increased sophistication of grouping strategies employed by the subjects over age levels. This finding is consistent with the Vygotsky and Bruner positions on concept formation generally, in that the development of concepts about word meaning appears to move from a fairly personalized level to a more formal structural level. The analysis of grouping strategies confirmed the finding that students in the eighth grade were beginning to be capable of making formal identification of some kinds of meaning.

There was some evidence to suggest that greater changes occurred between grades 5 and 8 than between grades 8 and 11. This is consistent with Piaget's (1950) view of cognitive development. Between grades 5 and 8 most of the children would be moving into the formal operations period which would allow them to attend to the structural and logical aspects of the definitions.

The evidence concerning the age at which the subject was able to classify the definitions may have been clouded by the high imagery values of many of the words used. Paivo (1969) states that there are two processes that underlie meaning, one deriving meaning from the interassociative relations among the words; the other deriving meaning from the imaginal representation of the words or units. Presumably if the image evoking qualities of the definitions used were high, it would be much harder for the subjects to attend to the structural qualities of the definition.

At a simple level, the evidence from the present study contributes to the utility of Evanechko's (1970) description of semantic space. At a more general level, the results may be useful for achievement test construction. Anderson (1972), in a discussion of the problems inherent in the construction of achievement tests to measure comprehension, points out that a method for deriving test items from instructional statements is needed, such that only those comprehending the statements can answer the items. Anderson suggests several ways in which comprehension questions can be formulated from transcripts of instruction. For example, questions may be formed by substituting synonyms or superordinate terms for words used in the instruction. His view is that to measure comprehension, items must be written which "occupy as much semantic space as possible."

The results of the present study suggest that for adolescent students, it may be possible to develop comprehension items by exploring the kinds of meaning listed in Table 1. For example, following a unit on statistics, the comprehension of the concept correlation might be explored by developing questions in which the student is asked to choose the statement which best illustrates the kind of meaning given. Definitions of the meaning categories would be supplied to the student. Five examples using only two alternatives are shown below. (The keyed response is marked with an asterisk.)

- | | |
|-------------------|--|
| 1. Similarity: | covariance—correlation* |
| | mean—correlation |
| 2. Superordinate: | statistic—correlation* |
| | standard deviation—correlation |
| 3. Coordinate: | correlation—mean* |
| | correlation—statistic |
| 4. Repetition: | correlation—correlation of a variable |
| | correlation—correlation between two variables* |
| 5. Use of: | correlation—for describing the relationship* |
| | correlation—for describing the variation |

In this way, items occupying a substantial portion of the semantic space could be developed.

References

- Al-Issa, I. The development of word definition in children. *The Journal of Genetic Psychology*, 1969, 114, 25-28.
- Anderson, R. C. How to construct achievement tests to assess comprehension. *Review of Educational Research*, 1972, 42, 145-170.
- Annett, M. The classification of instances of four common class concepts by children and adults. *The British Journal of Educational Psychology*, 1959, 29, 223-237.
- Bruner, J. S. The course of cognitive growth. *American Psychologist*, 1964, 19, 1-15.
- Bruner, J. S., & Oliver, R. R. Development of equivalence transformations in children. *Monographs of the Society for Research in Child Development*, 1963, 28, (6, whole No. 86).
- Burns, D. G. A note on the responses made by secondary school children in their definitions of words. *The British Journal of Educational Psychology*, 1960, 30, 30-39.
- Cronbach, L. J. Measuring knowledge of precise word meaning. *Journal of Educational Research*, 1943, 36, 528-534.
- Dale, E., Eicholz, G., & Bennet, B. *Children's knowledge of words*. Columbus Bureau of Educational Research and Service, Ohio State University, 1960.
- Evanechko, P. O. The dimensions of children's meaning space. Unpublished doctoral dissertation, University of Alberta, 1970.
- Evanechko, P. O., & Maguire, T. O. The dimensions of children's meaning space. *American Educational Research Journal*, 1972, 9, 507-523.
- Evans, G. T. The analysis of categorizing behavior. *Psychometrika*, 1970, 35, 367-392.
- Fiefel, H., & Lorge, I. Qualitative differences in the vocabulary responses of children. *The Journal of Educational Psychology*, 1950, 41, 1-18.
- Flavell, J. A., & Flavell, E. R. One determinant of judged semantic and associative connection between words. *Journal of Experimental Psychology*, 1959, 58, 159-165.
- Lewinski, R. J. Vocabulary and mental development: A quantitative investigation and review of research. *Journal of Genetic Psychology*, 1948, 72, 247-281.

The Development of Word Meaning Discrimination in Children

- Osgood, C. E., Suci, G. J., & Tannenbaun, P. H. *The measurement of meaning*. Urbana: University of Illinois Press, 1957.
- Paivo, A. Mental imagery in associated learning and memory. *Psychological Review*, 1969, 76, 241-263.
- Patsula, R. B. Children's categorizations of word definitions. Unpublished masters thesis. University of Alberta, 1972.
- Petty, W. T., Herold, C. P., & Stoll, E. The state of knowledge about the teaching of vocabulary. Report of Cooperative Research Project 3128. Champaign, Illinois: National Council of Teachers of English, 1968.
- Piaget, J. *The psychology of intelligence*. London: Routledge and Kegan Paul, 1950.
- Russell, D. H. The dimensions of children's meaning vocabularies in grades four through twelve. *University of California Publications in Education*, Vol. 11, No. 5, Berkeley: University of California Press, 1954.
- Swartz, K., & Hall, A. E. Development of relational concepts and word definitions in children five through eleven. *Child Development*, 1972, 43, 239-244.
- Vinacke, W. E. The investigation of concept formation. *Psychology Bulletin*, 1951, 48, 1-30.
- Vygotsky, L. S. *Thought and language*. (Edited and translated by Eugenia Haufmann and Gertrude Vakar.) New York: Wiley, 1962.
- Welch, L. A preliminary investigation of some aspects of the hierarchical development of concepts. *Journal of Genetic Psychology*, 1940, 22, 359-378.
- Wiley, D. E. Latent partition analysis. *Psychometrika*, 1967, 32, 183-193.
- Wolman, R. N., & Barker, E. N. A developmental study of word definition. *Journal of Genetic Psychology*, 1965, 107, 159-166.

B. Y. CARD

The University of Alberta

Community Vocational Centres in Northwestern Alberta

*A case study in social invention and development
in the field of adult education*

A Community Vocational Centre is a locally-managed adult education enterprise invented to meet primarily the needs of native Canadians in Alberta's isolated communities. Factors contributing to the invention included situational needs, new opportunities for occupations, new resources for educational upgrading, a climate of increasing aspirations and political awareness among native Canadians, and above all, strategically located and qualified personnel to undertake the inventing. Important organizational elements in CVCs include the local management committee, the educational technician, a regional vocational centre for training the technicians, a regional supervisory staff, flexible budgeting, and regional input into provincial policy making. After five years of steady growth in the number of CVCs and clientele served, educational attainment and aspiration among adults of the region's isolated communities has risen. The Educational Technician role has proven a significant channel of occupational mobility. However, as an innovation within the unpreempted social space of government administration and public education, CVCs still face critical challenges in local communities, in the region and in provincial decision-making arenas. (Dr. Card is Professor of Sociology of Education in the Department of Educational Foundations, University of Alberta.)

What are Community Vocational Centres?

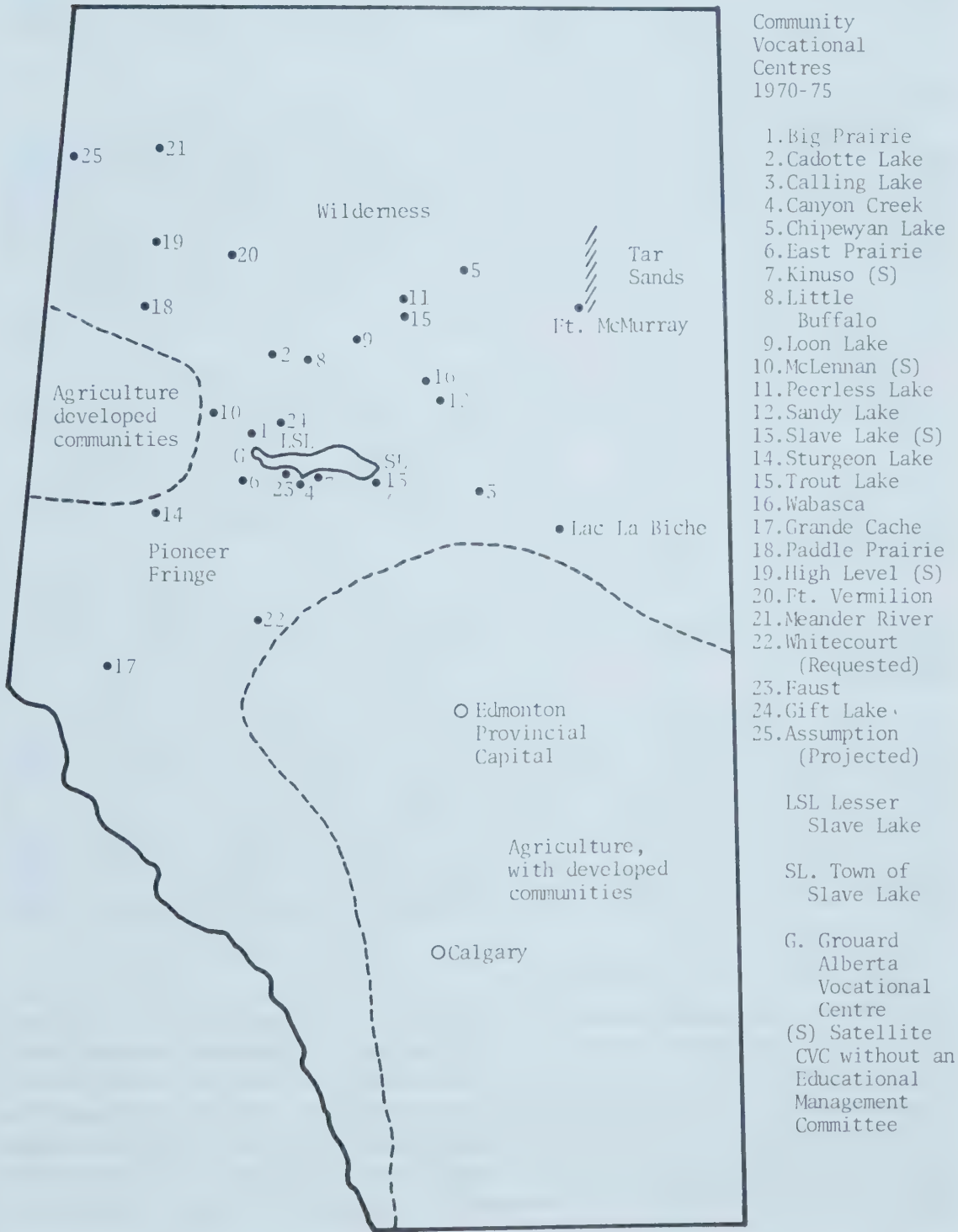
A Community Vocational Centre (CVC) is the name typically given to a locally-managed adult education enterprise in an isolated or frontier Alberta community where a large proportion of the population are of native Canadian ancestry. In its organizational and operational aspects a CVC is a group of local adults meeting at times appropriate for them in some premise in their community known as "The Centre." Here they are aided in learning by one or more Education Technicians, who know the native language and local conditions. The Education Technicians are often recruited, and in most cases

Community Vocational Centres in Northwestern Alberta

selected, by the local Education Management Committee, which is responsible for the CVC. This committee, in addition to selecting the education technicians, is also concerned with recruiting and selecting students, operating the Centre premises, deciding on the programs offered, requesting enrichment and support services from outside the community, and negotiating and spending a budget. This is the CVC as a local enterprise. Their geographic distribution is shown in Figure 1.

FIGURE 1

SCHEMATIC MAP OF ALBERTA SHOWING COMMUNITY VOCATIONAL CENTRES IN NORTHWESTERN SERVICE AREA OF DEPARTMENT OF ADVANCED EDUCATION



However, such local enterprises require support from a regional level, especially when scattered over an area of 93,000 square miles, much of which is primal wilderness. Technical support is supplied from the Grouard Alberta Vocational Centre on the west end of Lesser Slave Lake. The Grouard AVC trains Educational Technicians and helps prepare their instructional materials for CVCs. Administrative support is provided for by a small staff of bush-experienced educator/community development personnel attached to the regional office of the Alberta Department of Advanced Education at the town of Slave Lake, on the east end of Lesser Slave Lake. This staff is responsible for coordinating, supervising, developing, promoting, evaluating CVCs, and negotiating budgetary requirements with the Department of Advanced Education. Between 1970 and 1975 a total of 23 CVCs were or had been in operation and two were about to be established.

Background, Origins and Objectives

The first CVCs were created in 1970. They were a relatively complex social invention which reflected the operation of three sets of factors, those reflecting specific needs of Northwestern Alberta, those reflecting new opportunities and resources available for the region, and most important, the availability of a pool of concerned and talented persons who were free to undertake inventing responsibilities. The interplay of these factors is depicted schematically in Figure 2. Eleven factors identifiable in the inventing process are as follows:

Situational-need Factors

1. Community change and trauma from successive Mainstream Society penetrations of the area, with a consequent lessening of the will and the ability of local inhabitants to gain a living from hunting, fishing and trapping.
2. Commercial exploitation and depletion of lake and forest resources, especially around Lesser Slave Lake, which jeopardized seasonal work opportunities for a rising population.
3. Virtual entrapment of some native Canadian segments of the population in lower social strata characterized by poverty, high alcoholism and crime rates, unemployment, and anomie. (Card, Hirabayashi, & French, et al., 1963; Hatt, 1967.)
4. Low educational attainment of the region's population generally, and more especially of the persons of native ancestry among whom there was an excessively high rate of youthful dropouts and a significantly high rate of adult illiteracy. (L. W. Downey Research Associates, 1974, pp. 57-60.)
5. The expanding role of government in the region, which, while extending controls, service and economic resources from Mainstream Society to the area, also tended to decrease the sense of independence and the decision-making power of local residents.

Opportunity-resource Factors

6. The assault on the problems of the region by a succession of researchers, community development workers and strategists bringing new knowledge, increased motivation and new mobilizations of resources to bear on these problems. Their efforts led federal/provincial governments to create the Lesser Slave Special Area for regional development in 1970, with a five year funding from the Department of Regional Economic Expansion. (See Appendix I for a list of influential documents.)

7. The availability of personnel, educational technology and educational programs from Alberta and Saskatchewan Newstart pilot projects for educating native Canadian adults.

8. The beginning of commercial exploitation of the Athabasca tar sands in Northeastern Alberta and the anticipated demand for trained workers of this and future tar sands developments.

9. The Grouard Vocational High School, opened in 1963 for the youth of the region, especially those of native ancestry (Chalmers, 1967, pp. 217-8), was closed in 1968. This closed, well equipped facility, complete with residences, was a source of embarrassment to education officials and a focus of native Canadian lobbying for their adult education.

10. The emergence of native power and influence as organizations were formed and as leaders emerged among both Indians and Metis of the region.

11. The presence and availability in the region of senior education officials concerned with community-level and regional vocational education, of personnel with expertise and experience acquired in Alberta Newstart, and of community developers from the Lesser Slave Lake Task Force of the Alberta Human Resources Development Authority, who were in constant dialogue with residents of the region, particularly residents of isolated communities.

The inventing of the CVCs occurred in the late 1960s when Alberta was developing a three-tiered province-wide system of vocational education:

1. A provincial-level set of technical institutes augmented by universities.
2. A regional-level set of vocational centres for adults and augmented by regional colleges.
3. A community-level set of vocational high schools, adult education programs and apprenticeships, which would meet local employment needs, facilitate entry of local persons to agencies offering advanced training, and contribute to the mobility of the labor force in relation to work opportunities in the province.

CVCs were conceived by both inventors and provincial policy-makers as community-level agencies uniquely appropriate for underdeveloped or isolated communities with native Canadian populations.

The objectives of CVCs were viewed operationally within the larger framework of regional development. According to HRDA's 1971-72 Program A they were to (1) provide introductory courses in occupationally oriented skills, (2) provide training close to home community, (3) provide foundation for additional training opportunities for employment or advancement because of training, (4) improve leadership activities and skills, (5) develop individual talents, (6) raise aspiration levels.

In addition to these objectives Villett (1969, pp. 3, 4) saw CVCs helping adults acquire enough basic skills to participate adequately in the English-speaking community, upgrading youthful academic dropouts, and helping local communities create an advantageous and supporting climate for children to continue education in local schools.

The CVCs, as a requirement for obtaining designated-area funds, were also expected to serve as vocational information centres for the community. Underlying all these objectives was an unwritten goal of maintaining and increasing the dignity, respect and self-worth of native Canadians by recognizing

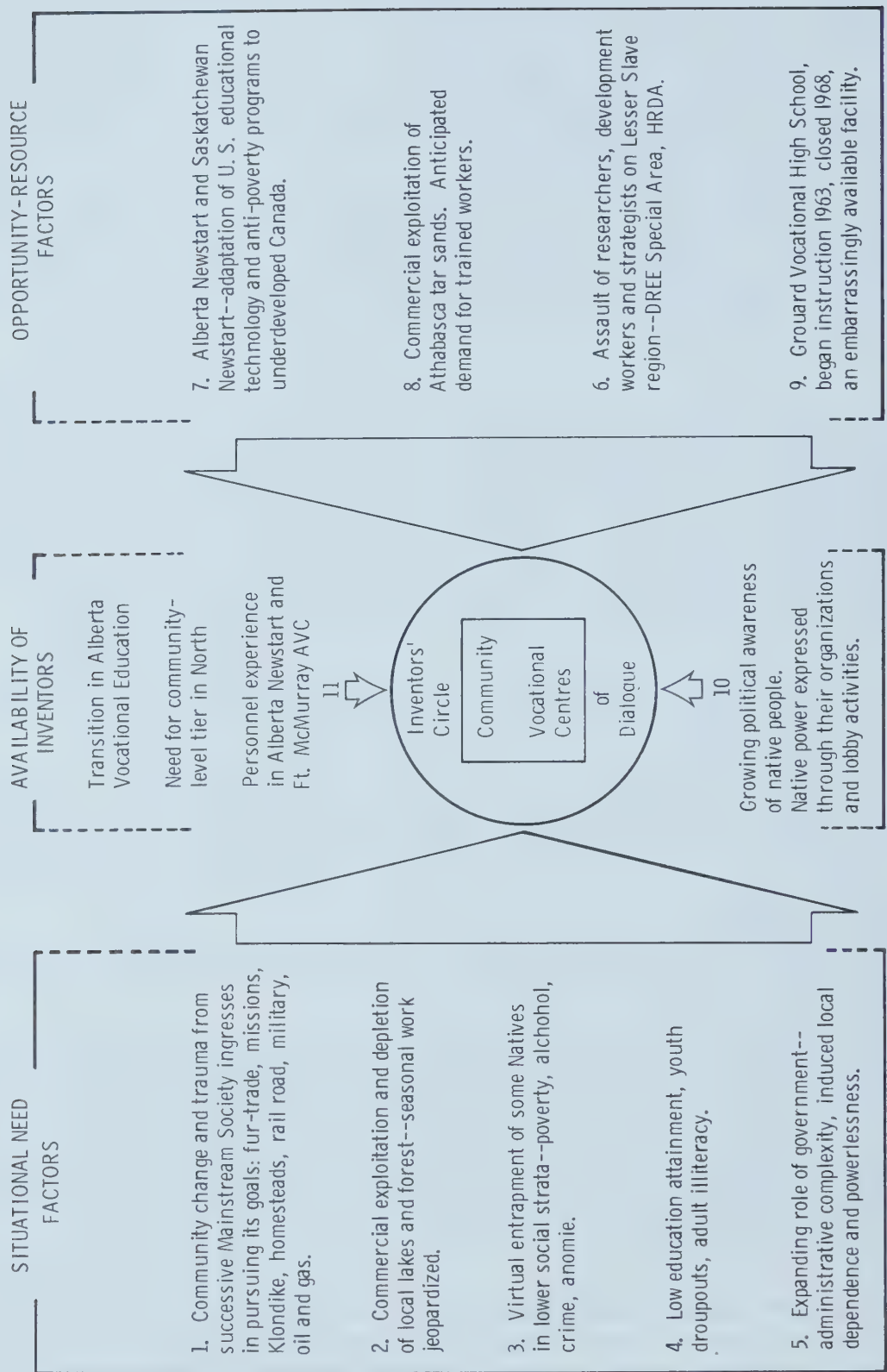


FIGURE 2
SCHEMATIC PRESENTATION OF FACTORS CONTRIBUTING TO THE
INVENTING OF COMMUNITY VOCATIONAL CENTRES IN
NORTHWESTERN ALBERTA 1967-1970
(Numbers refer to description of factors in the text)

ing the importance of their language and culture in CVC operations and by sharing and honouring the decision-making power of the locally-elected Educational Management Committee responsible for the CVC in its own community.

These early goals have remained constant in the operation of CVCs in isolated communities. As satellite CVCs have been established in larger, incorporated communities as a requested service operated by the Regional Office of Advanced Education, the goal of community involvement and management has been more difficult to achieve. (See Figure 1 for location and Figure 3 for organization of satellite CVCs and self-managed CVCs.) The organizational, governing structures and funding arrangements for achieving these goals are next considered.

Organization and Governing Structures

The full range of organization and governing structures for CVCs is shown in Figure 3. At the local level a self-managed CVC has three organizational components.

1. *The Educational Management Committee or EMC.* The EMC consists of three to five democratically selected persons, usually Cree-speaking in isolated communities, who meet as their affairs require. They assess and prioritize community needs for adult education and leadership development. They also request and administer adult education programmes, the latter including student selection and termination, acquiring and maintaining facilities, hiring and supervising Educational Technicians for upgrading programs and staff for short courses, and evaluating outcomes. They are recognized by the Minister of Advanced Education as having a legitimate involvement in all aspects of their local adult education. They may claim \$15 for each EMC evening meeting, \$25 for each day-long meeting as individual honoraria, and travel expenses for required meetings away from the community. Satellite CVCs, operated directly by the coordinator of CVCs as a service in larger, incorporated communities, make use of local existing organizations for advice, and do not have an EMC as such.

2. *The Instructional Staff.* The key instructor is the Education Technician engaged to conduct academic upgrading courses. The Education Technician has been trained at Grouard for providing "Individualized Programmed Instruction" to illiterate adults (grade 0) and all other grades up to and including 9. In six instances, where the Education Technician also has teacher certification, instruction is given for grades 10-12. The training of an Education Technician is shown in Figure 4. The pay rates are from \$668 to \$855 per month, depending on experience and qualifications, rates which are higher than for teacher aides employed in Northland schools, but lower than for professional certified teachers. Under the supervision of the EMC, the Education Technician takes full charge of CVC up-grading instruction. Their informal roles often include some counselling, recreation leadership, and community development. They frequently recommend their advanced students for Education Technician training at Grouard, and themselves often move to other work or advanced training. They tend to be a flexible and upwardly mobile component of the CVC organization and in this sense serve as dynamic role models for the community.

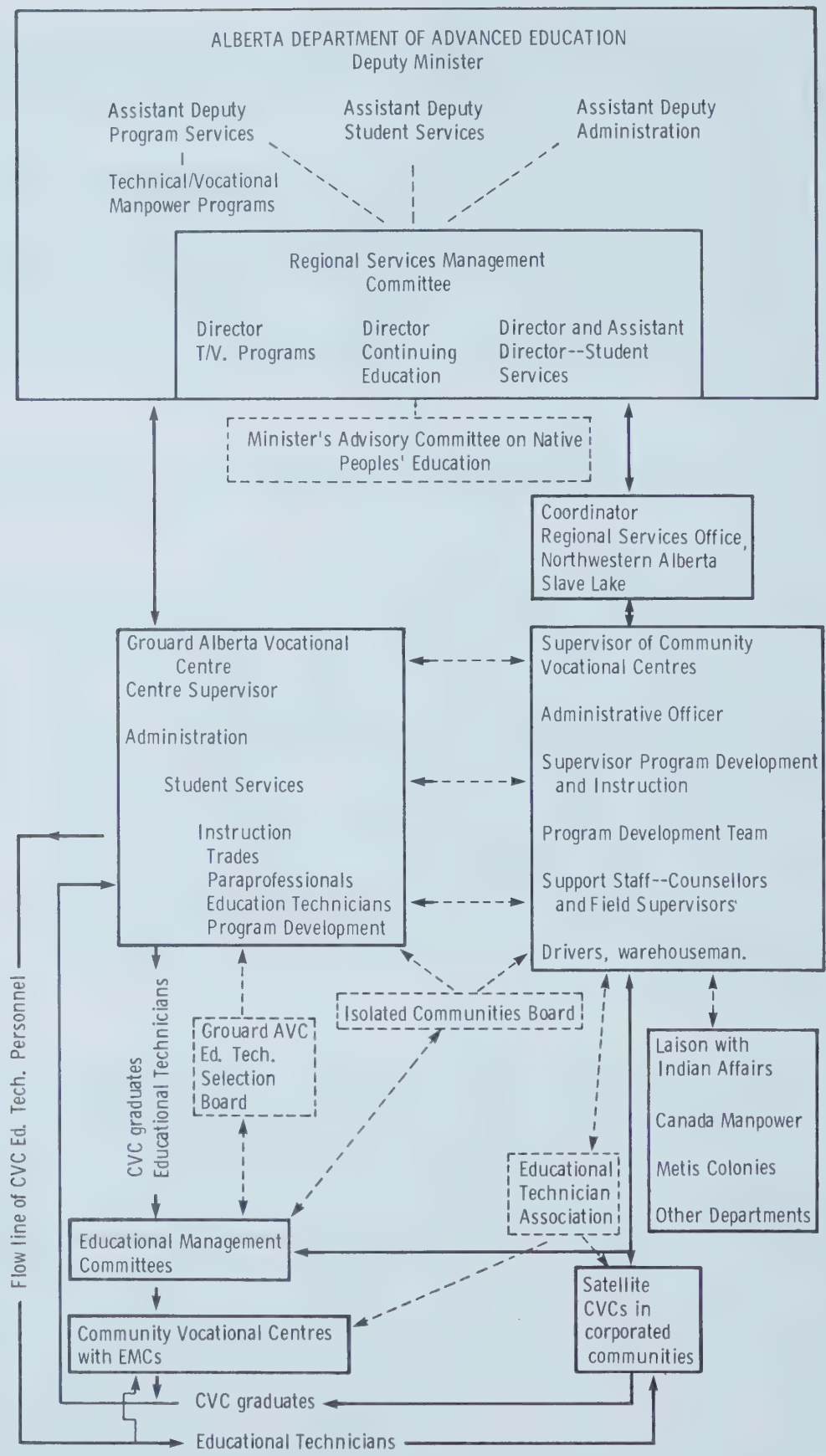


FIGURE 3
ORGANIZATION AND GOVERNANCE OF COMMUNITY VOCATIONAL CENTRES

Lines of control _____ of advice - - - -

Other instructors are brought to the community by EMCs for short courses or workshops in such subjects as house wiring, cooking, or leadership development. Such instructors and necessary equipment are secured by the EMC on a contract basis.

FIGURE 4
THE TRAINING OF AN EDUCATIONAL TECHNICIAN AT GROUARD
ALBERTA VOCATIONAL CENTRE, 1974-75
(Source: D'Aoust, 1974)

DURATION AND SEQUENCE OF EXPERIENCES :

Eight months: six months spent at Grouard AVC first, usually followed by paid teaching in the field, followed by two months back at Grouard during the next summer.

COURSE CONTENT

A. Teaching Materials Covered:

0-4 Mathematics	5-9 Mathematics
0-4 English	5-9 English
0-4 Reading	5-12 Reading
	5-9 Science

B. Practice Teaching -- 5 weeks.

C. Skills Training:

Audio-visual	25 hours
Typing	50 hours
Community and Leadership Development	30 hours
Public Speaking	30 hours
First Aid	Time as needed
Practicum in Speaking and Interpreting Native Languages	Time as needed
Life Skills Training	Time as needed

3. *CVC Clientele.* Adults and the older youths who have been out of school a year make up the clientele of CVCs. Those who are admitted as full-time trainees for up-grading are eligible for training allowances from Canada Manpower, Indian Affairs, or if they do not come in under these auspices, from Advanced Education. The 1974-75 allowance scale was \$6 per day for a single person living at home to \$18 per day for a person with four or more dependents. Since CVCs offer individualized instruction on a continuous intake and exit basis, trainees may enter or leave their program at any time for family or work

reasons, and when they are again free, carry on from where they left off. Students taking short courses do not qualify for training allowances.

As shown in figure 3, there are two major structures serving CVCs at the regional level. The Grouard AVC prepares Education Technicians as part of providing basic adult and vocational education for the region as a whole. The AVC is headed by a Centre Supervisor who reports directly to the Department of Advanced Education at Edmonton. The AVC instructors, counsellors and program developers preparing Education Technicians are in contact with CVCs through consultation with CVC supervisory staff at Slave Lake, through the Education Technician Selection Board made up of EMC representatives, and through association with Educational Management Committees from the entire region who hold meetings up to four times a year at the AVC.

The CVC supervisory staff, stationed at Slave Lake Regional Office of Advanced Education, is headed by a Supervisor of CVCs, who is directly responsible to the Regional Coordinator of Advanced Education. The Supervisor of CVCs is assisted by an Administrative Officer, a Program Development Coordinator and Academic Supervisor. Each of these three persons has a lengthy background of experience in northern Alberta communities. They are assisted by a support staff of Cree-speaking counsellors and field supervisors, most of whom have been Education Technicians, and a truck driver and warehouseman to look after supplies. In addition the Supervisor of CVCs is responsible for a Program Development Staff of three, stationed at Grouard where they are engaged, in collaboration with Grouard AVC program developers, in preparing instructional materials for CVCs.

As seen in Figure 3, there are two advisory bodies at the regional level, the Isolated Communities Board, which draws some members from local Educational Management Committees, and an association of Education Technicians which in 1974-75 was in the beginning stages of organization. At the provincial level the Minister of Advanced Education has an advisory committee on native education. At both regional and provincial levels, liaison is maintained with government departments, Canada Manpower, Indian Affairs, and Metis colony administration in CVC policy-making, program development, and the funding of trainees.

As the name implies, the important operations and involvements in the long run are at the community level, not as much the regional or provincial levels, except as these, through policy, staffing and budgetary arrangements, facilitate grass-roots community interaction and learning. The pattern for this was established by HRDA, and has been continued through the CVC Supervisory Staff at Slave Lake. Budgets are worked out in consultation with EMCs, on the basis of what they expect to do—enroll so many adults for upgrading, pay so much in training allowances, put on so many short courses, etc. These individual community budgets are aggregated by the Supervisory Staff and forwarded to the Department of Advanced Education for approval and consolidation into the budget submitted to the provincial legislature. This process has been likened to a private enterprise component in government, with each EMC in the role of a local entrepreneur in adult education. A diverging view is held by some middle-level government officials, who insist on budgeting not as community involvement but as the necessary preparation for the delivery of another government service. Given the power of an EMC to select or dismiss an Education Technician, to admit and terminate studentship, to contract for

short courses, and to control and manage local facilities, their budgeting is part of community involvement. They are in a sense the community's executive committee for adult education. In larger communities this degree of community control has not been evolved, though existing community organizations participate, somewhat intermittently, in a planning and advisory role.

As for actual budgets, the total estimated budget for CVCs in 1973-74 was \$408,531 for tuition costs, \$156,992 for training allowances.² This was for an estimated 300 trainees, for whom per capita costs would be \$1,361 for tuition, \$523 for training allowance. The 1974-75 estimates were \$493,100 for tuition, \$381,077 for training allowance for an estimated 390 trainees, a per capita rate of \$1,264 for tuition and \$977 for allowance. The 1973 per diem cost of training, including allowances and tuition, is estimated at \$9.00, while in 1974-75 it approximates \$13. These costs compare favorably with costs in post-secondary or adult-education institutions of Northwestern Alberta and adjoining British Columbia, where, excluding CVCs, they range from \$8.27 to \$10.35 for instruction only, and from \$14.67 to \$31.71 for total operating costs, per student day. (L. W. Downey Research Associates, 1974, p. 89). However, these comparisons should not obscure the differences in budgets for individual communities, whose volume of potential trainees and their economic need can be taken into account by the EMCs. The principle involved here is that of locally-initiated unequal funding to provide equality of opportunity for persons whose ethnicity, language, age, location, socioeconomic status, and traditional occupations for many years have rendered other approaches to educational equality ineffective and frustrating for all concerned.

Current Status of CVCs and Observable Results

Since their inception in 1970-71, CVCs have increased steadily in numbers, in the number of students for which they have training places, and in the number of persons trained. In 1971-72, 175 students received approximately 8,000 days of training (Department of Advanced Education, 1972, p. 9) in eight CVCs. In 1974-75, it was estimated that 390 students would receive 67,250 days of training. During the mid-year there were indications that as many as 500 adults would be receiving training in 0-12 type basic and vocational courses and in short courses. While firm statistics were not available, supervisory staff estimated that educational and vocational aspirations of the student clientele were changing from basic literacy to completion of Grade IX as an expectation. In 1974-75 approximately 73 per cent of the students were at or near the 9 level, 40 students viewed 10 as their level; 14, grade 11; and 9, grade 12. Further, 15 per cent of the students, some aspiring to occupations requiring high school, were taking work at the 10-12 level. This is in a region where in 1961 one-third of all native adults had no education and only 0.8 per cent Grade 10 or more (L. W. Downey Research Associates, 1974, p. 58).

Perhaps the most remarkable transition is observable among Education Technicians. Between 1970 and October 1974 (D'Aoust, 1974) Grouard AVC had graduated 81 and had career information on 70. Of the 70, 29 were working at the AVC or in CVCs. Nine were working in related fields, but not with the Department of Advanced Education. Eight were working in unrelated fields, for example, laborers, while 8 had completed one or more years of university but had interrupted studies for a year to earn money, 7 were attending the University of Calgary, 4 the University of Alberta at Edmonton, 2 the University

of Saskatchewan, 1 was taking teacher training in the Northwest Territories, 1 had become a journeyman welder, while another was apprenticing in pipe trades at AVC Ft. McMurray. What the CVCs appear to have done, in addition to upgrading adults and raising their aspiration levels, is to have created a flexible job category, between that of an "aide" and a professional teacher, which has begun to meet mobility needs of an important segment of native young adults.

Two other results have become apparent, both not without risks to the future of CVCs, especially if they should expand rapidly. The first is that the growth in leadership capacity of EMC members is highly observable but also uneven, with the result that those manifesting leadership ability are frequently overworked locally or lured away from EMCs by requests for service at the regional level. Nevertheless, the pool of talented local leaders has increased through the work of the EMCs. The second is the growing demand of non-isolated communities and of different agencies for a CVC or a very similar organization, staffed with Education Technicians or their equivalent in some other field. Requests for "satellite" CVCs or for out-reach CVCs for persons in jails have increased and some have been set up, though when this happens the local control aspect tends to diminish.

A final accomplishment should not be overlooked. In six communities CVCs have built their own buildings, which have been sold back to the communities for the nominal charge of one dollar. These building projects have had an educative value for the entire community. It has been noted that vandalism has been absent where the community has done the building and taken over control of the CVC.

Difficulties, Problems and Crises of CVCs as Innovations

CVCs, beset with a bewildering array of problems locally, regionally, provincially and even nationally, have emerged from their formative years as adaptive, but still fragile outposts of democracy in adult education in a dynamic frontier situation. Their future appears to rely on two major factors, the strength of basic democratic process in the communities of the region, and a willingness of other levels of government and Mainstream Society to tolerate, facilitate and support this process. The array of difficulties and problems is formidable and crises of various degrees are virtually a daily or weekly occurrence. The following listing attempts to identify some of the major challenges of the CVCs.

Locally:

1. Excessive drinking, quarrelling, and outbursts of hostility and violence tend to paralyze and slow down community action in a number of the communities with CVCs, particularly along Lesser Slave Lake. The intense social problems of some communities tend to distort the goals and undermine the usefulness of any form of adult education and to make force appear as the most expedient solution for community problems.

2. Education is still not a priority value and needs development, though the situation has changed for the better as CVCs have begun to give new accessibility and meaning to education locally.

3. Overwhelming of the few willing, capable leaders with community responsibilities, to the point they have difficulty earning a living or caring for their families, especially as trappers or seasonal workers.

4. Abdication of EMC power and responsibilities by local EMC members to an Education Technician.

5. The lack of programmed instructional material for grades 10-12 puts an excessive load on the Education Technician. CVCs were originally designed and equipped with programs for 0-9. High school adaptations have been slow, partly because of divided opinion about their usefulness and legitimacy so far as CVCs were concerned.

6. A feeling among locals that much of the programmed instruction could be better related to their day-to-day concerns and experience. Life skills remain somewhat unintegrated with academic up-grading.

7. Some Northland School youth appear attracted to the idea of education with pay, so drop out and wait until they can enter a CVC, where they learn faster than in school and have an income. The actual numbers who do so are very small.

8. Persons without other means of support, face hardship from delays in processing training allowance applications through Edmonton computers. This has been overcome by arrangements for emergency, short-term help from the region's Social Development office. The problem of transferring medical coverage from persons transferring from Social Development assistance to educational training allowances has been overcome.

9. The intrusion of provincially complex government services into small, isolated communities overloads the community's committee-carrying capacity. The EMC may be competing for local time and talent with an Early Childhood Services Committee and Northland Division local school advisory councils, where they exist, with a committee for Preventive Social Services projects, another for the school's hot lunch program, with a local association required for the public nurse, with a recreation committee, or with a Local Further Education Council of Advanced Education, all of which are needed to have access to certain government funds or services. Top this mélange with a local Metis Association, an Indian Band Council, and one or two church-related voluntary associations, and the small local community even in the Northwestern Alberta becomes a complex and competitive environment for adult education.

10. CVCs have not resolved the problem of culture shock as local trainees migrate to larger centres for further training or work.

Regionally:

1. Sheer distance between communities is a major problem. This has been alleviated in part by new road construction. CVC supervisory staff resolve it their own way, by being private pilots.

2. Rival views about whether CVCs should be outreaches of an institution such as Grouard AVC or independent agencies supervised by persons with a community-development background and interest. The basic issue here is related to programs and their administration on the one hand, versus process, with the programs to express and aid process on the other. Grouard and Slave Lake appear to have worked out an accommodative division of labor here that is difficult to transmit to the uninitiated.

3. Rapid expansion of CVCs, especially in larger communities of the regions, where there has not been the time nor the energy put into developing

process aspects because of the immediacy of problems or the strong demand for academic up-grading programs by other funding agencies, especially on the part of Canada Manpower.

4. Modifications to the Education Technician role from those regional departments who would adapt it for social work, rehabilitation of prisoners, health education, or recreational leadership. Their single or combined efforts could render the new Technician a very powerful community figure, but without a local group to provide legitimation, supervision and appeal from his power. The EMC serves this purpose for the Education Technician who currently has a livable, useful, but non-threatening community role.

5. Expansion of CVCs outside the Northwest Alberta initial "Designated Area" puts strain on the capacity of Grouard AVC to train Education Technicians for a range of community environments and on the ability of the CVC Supervisory Staff to give sufficient attention to the local management component.

Provincially:

1. The organizational positions intervening between grass-roots EMCs and top provincial policy makers are filled with many highly-trained and dedicated persons recruited from developed Central and Southern Alberta concerned primarily with implementing policy in a regulated manner. The private enterprise, developmental and process aspects of CVCs do not conform to their traditional expectations. Recurrent instances of misunderstanding have to be worked out.

2. Major policy research in Alberta education has paid very little attention to CVCs. The Grouard AVC has been "phased out" of its contributions to CVCs in one long-range forecast (Commission on Educational planning, 1972, pp. 93-94), while another (L. W. Downey Research Associates, 1974, pp. 30-34) puts the Grouard centre on trial for two years. These studies imply a problematic future for CVCs, while at the same time providing no evidence that their unique contributions have been considered seriously.

3. The Education Technician role could readily lose its legitimacy and its flexibility from two sources: (1) regulations regarding teacher certification which have been carefully guarded by both the Alberta Department of Education and the Alberta Teachers' Association on the one hand, and (2) provincial civil service on the other, which could provide a permanent "slot" for Education Technicians fixing their classification, job descriptions, pay rates, retirement, and other matters, at a time when role flexibility has been and continues to be a major asset for the Education Technician.

4. Provincially-sponsored formula financing or its equivalent, a means used to equalize opportunity on a mass scale, could undermine the adaptability of CVCs with their locally initiated and managed budgets. As long as funding can be kept flexible and adequate, the CVCs stand a chance of surviving.

5. The Indian Association of Alberta has had, in conjunction at times with the Alberta Metis Association, designs for a native education system of education. Should such a system emerge, the CVCs as now organized and funded, could undergo modification or be discarded for some other educational arrangement, though CVC principles of operation might be retained for their essential usefulness.

6. The attraction and retraction of federal funding could change the CVCs. For example, a reclassification of Education Technicians to "Teacher Assistants" would allegedly open up grant possibilities from some federal programs. So far CVCs have withstood retraction of federal funds as the 1970 federal/provincial Special Area agreement ended in 1975.

7. A major threat at the provincial level is for a change of government to abolish or alter CVCs as a matter of take-over policy. The CVCs have passed through one change of government and survived, though the creative, facilitating arrangement known as the Lesser Slave Lake Task Force of HRDA did not. Perhaps the lesson here is that it is politically easier to alter regional organization than something so fundamental at the grass-roots level as a CVC. An alternative view might be that the CVCs persisted because they were perceived as a useful social invention by informed persons in all levels of government.

Underlying all the above problems is the possibility that CVCs, by successfully up-grading the educational levels and vocational skills of local community residents, may no longer be needed as they currently exist. When this condition arrives, new social inventions serving Northwestern Alberta communities will be required, but these can be developed from a much better basis of knowledge and experience because of CVCs.

References

- Alberta Human Resources Development Authority, Research and Planning Division. Lesser Slave Lake special area proposed programs 1971-72. Mimeographed. 1971-72.
- Chalmers, J. W. *Schools of the Foothills Province: The story of public education in Alberta*. Toronto: The University of Toronto Press, 1967.
- Commission on Educational Planning. *A future of choices: A choice of futures*. (The Worth Report) Edmonton: Queen's Printer, 1972.
- D'Aoust, B. (Mimeographed recruitment letter to potential Educational Technicians, with enclosures.) Grouard: Alberta Vocational Centre, 1974, October 15.
- Department of Advanced Education. Annual Report. Edmonton: Queen's Printer, 1972.
- L. W. Downey Research Associates. *Toward a system of post secondary and adult education for Northwestern Alberta and Northeastern British Columbia. A report to the governments of Alberta and British Columbia*. Edmonton: L. W. Downey Research Associates, 1974.
- Villett, L. S. Submission to the Worth Commission concerning adult vocational training from the viewpoint of the Vocational Training Section of the Division of Vocational Education, Department of Education. Edmonton: Mimeographed, 1969.

Presented at the Annual Meeting of the Canadian Sociology and Anthropology Association, Edmonton, May 30, 1975. This study was commissioned as part of a nation-wide survey of Canadian development projects potentially of interest to developing countries. The survey, funded through the Canadian International Development Research Centre and the Ministries of Regional Economic Expansion and of the Secretary of State, is directed by Professor Glen Eyford, Director of the M.A. Program in Community Development, University of Alberta.

¹ The information on which this report is based was obtained through observation, interviews and documentary study. Where data are taken directly from published sources, these have been cited. The writer takes responsibility for errors, omissions and interpretations presented here. At the same time he is mindful of his debt to those who so generously shared their experience and information related to CVCs. This report does not deal with the outreach CVCs operated in Northeastern Alberta as satellites of Lac La Biche Alberta Vocational Centre.

² Calculations by the writer, based on CVC appropriations in Appropriation Nos. 3063 and 3075, Department of Advanced Education, years ending March 31, 1974 and 1975.

APPENDIX

A Chronological List of Influential Reports and Proposals Relating to Northwestern Alberta and Native Canadians in Northern Alberta, 1958 to 1972.

Royal Commission on the Development of Northern Alberta

1958 Report, Edmonton: Queen's Printer.

Card, B. Y., G. K. Hirabayashi and C. L. French, in collaboration with S. Greenhill, B. A. Ruether and R. S. MacArthur

1963 The Metis in Alberta Society, with Special Reference to Social, Economic and Cultural Factors Associated with Persistently High Tuberculosis Incidence. Report on Project A(1960-63) University of Alberta Committee for Social Research, Prepared for The Alberta Tuberculosis Association. Edmonton: The University of Alberta.

Hatt, Fred K.

1967 Community Opportunity Assessment: Appendix A—Metis of the Lac La Biche Area. Edmonton: Human Resources Research and Development, Executive Council, Government of Alberta.

Banta, Gordon

1968 A survey of Problems in the Lesser Slave Lake Area. A Report to the Human Resources Development Authority, Government of Alberta. (Mimeographed.)

Bigam, Jerry F.

1968 Industry and Resources in Alberta's Census Division 15. Edmonton: Rural Development Research Branch, Economics Division, Alberta Department of Agriculture.

Rural Development Research Branch, Alberta Department of Agriculture

1968 An Analysis of Resources in Alberta's Lesser Slave Lake Area.

Alberta Human Resources Development Authority Research and Planning Division

1969 The B-15 Plan: An Outline for Rural Development in Alberta's Census Division 15. Edmonton: HRDA.

Government of Alberta

n.d. An address to the Honourable Jean Marchand, Minister of Regional Economic Expansion and Proposed Comprehensive Regional Development Plan for Lesser Slave Lake Area of Alberta, delivered by the Honourable Ray Speaker, Minister of Social Development, Chairman of the Human Resources Development Authority and the Honourable Ray Ratzlaff, Minister of Industry and Tourism, Chairman of the Physical Resource Development Committee of the Government of the Province of Alberta.

Department of Regional Economic Expansion

1970 Canada/Alberta Agreement in Special Area of Lesser Slave Lake 1970-72.

Aumack, Lewis

1972 Research Evaluations of Alberta Newstart Programs and Operations. Lac La Biche, Alberta. Alberta Newstart Inc.

J . P . D A S

The University of Alberta

J . M A N O S

Dalhousie University

R . N . K A N U N G O

McGill University

Performance of Canadian Native, Black and White Children on Some Cognitive and Personality Tests

Cognitive abilities and personality characteristics were examined in Grade 4 children from (i) low and high socioeconomic strata from the white population in Edmonton, (ii) white and black populations in Montreal, and (iii) native children from a reserve school near Edmonton. Parents of these children were also given some personality tests and were interviewed to obtain an index of educational environment in the child's home, reflecting parents' attitude to the child's education. The results indicated that Canadian native and white groups showed differences in their performance on some of the cognitive tests. Personality differences, as measured by Eysenck's extraversion-neuroticism questionnaire and questionnaires for rigidity and locus of control, were negligible between the groups of children at the Grade 4 level. However, the adults in the three ethnic and two socioeconomic groups showed some significant personality differences. Performance of the children on reasoning and school achievement tests were predictable by parental attitude to child's education. This finding was discussed in terms of the role of personological and social variables which may affect the development of a child's cognitive competence. (Dr. Das is Director of the Centre for the Study of Mental Retardation, Faculty of Education, The University of Alberta; J. Manos is Assistant Professor of Education, Dalhousie University; Dr. Kanungo is Professor in the School of Management, McGill University.)

In the study of cognitive competence, there seems to be a shift in emphasis from "abilities" to an inquiry into the "processes" which determine an individual's level of performance. The shift has been influenced partly by the need of comparing the same "ability" in different cultures (Cole & Bruner, 1971) and of explaining such differences. It has been argued that the researcher

should consider the lower level performance in certain cultures as a difference rather than a defect as compared to the standards of performance by groups in Western culture. The cognitive style or strategy—an habitual mode of processing information adopted by a group—may account for the lower performance of the disadvantaged child (Das, 1973a). For example, Jensen (1971) observed that children from the middle to upper socioeconomic classes use reasoning and abstraction much more than those from the lower socioeconomic classes. The latter are quite proficient in associative learning, however. If one considers abstraction and associative learning as processes, one should then investigate the conditions which cause one group to use one process, such as abstraction, and not the other for integrating information. Evidence is now available to suggest that a particular mode of information integration, such as successive processing, may be used by certain non-white groups in tasks which usually elicit simultaneous processing in white children (Das, 1973b). Discussing a study on learning, race and school success, Rohwer (1971) not only noted differences in learning styles between ethnic groups, but also could demonstrate substantive improvement in the performance of minority children by teaching them an appropriate learning strategy for the task at hand.

Perhaps a general orientation for comparing intellectual performance among culturally different groups can be proposed at this point. Following Luria (1971), one may view cognitive activities as “a social phenomenon in origin, and as processes formed during the course of mastery of general human experiences” (p. 272). In other words, intellectual processes are not unchangeable and universal, nor should they have an *a priori* characteristic as is very often reflected in IQ. They are shaped by the experience through which a subgroup passes, and these experiences, or “historical developments” as Luria calls them, are reflected not only in what we think, but also in the structure or style of our thinking. Cognitive processes “are not independent and unchanging ‘abilities’ or ‘functions’ of human consciousness; they are processes occurring in concrete, practical activities, and are formed within the limits of this activity” (p. 266). Cole and Bruner (1971) essentially make the same observation by reminding one “that the most important thing about any ‘underlying competence’ is the nature of the situations in which it expresses itself” (p. 784).

Although one cannot deny the influence of genetic factors in determining cognitive competence, we have chosen to accept the views of Cole, Bruner and Luria as a basic framework to understand the effects of socioeconomic status and personality on a group’s cognitive performance. In the present study, we have examined children from Grade 4 who belonged to white, black and native communities by giving them some cognitive and personality tests. The object is to find out how well the extra-intellectual factors, such as socioeconomic status, parental attitude towards the child’s education and belief in an internal vs. external locus of control for one’s behavior, can predict performance in some cognitive tests.

The effect of low socioeconomic status on school achievement is another concern of this study. It has been widely discussed before (Deutsch et al., 1967; Jensen, 1969). A review of the literature led Das (1973a) to conclude that the lower class child is more prone to academic failure than the middle class child. It may appear that the language of the former may stand as a major impediment in his scholastic progress. But the review suggested that it was not his

grammar which prevented him from doing well in school; rather, the total attitude of the lower class child and those of his parents or the community towards his education accounted for his poor performance. Thus, parental expectation regarding the child's education seems to be an important determinant of individual differences in school achievement.

In addition to expecting differences between the groups on cognitive tests in the present study, differences in some personality characteristics are anticipated. Both children and their parents would be examined for this purpose. It is hoped that because of the wide variations between the groups' socioeconomic class and/or ethnic membership personality differences, at least among the adults, are likely to occur. A person's class makes him conform to the image of the group (Sherif, 1947). However, this part of the study is of secondary importance, as only a few tests of personality which have known influence on learning and cognitive processes have been used in the present paper.

Method

Samples

The white children in Edmonton were taken from Grade 4 classes of public schools. The native children were from Grade 4 of a Cree Indian reservation school at Hobbema, about 40 miles from Edmonton. The white sample was divided between high (middle to upper) and low socioeconomic (SES) groups as determined by the occupation of the chief wage earner in the family on the Blishen scale (Blishen, Jones, Naegele, & Porter, 1965). The means and standard deviations (SDs) of the two white groups on this scale were 59.50 ± 8.34 for the high and 40.20 ± 4.60 for the low SES group. It is inappropriate to apply the occupational scale to the native parents, but they could be easily categorized as low SES in terms of income for comparison purposes with the white groups. The number of children in each group was 30.

The white and black children in Montreal were taken from Grade 4 of the city's schools. Their SES was determined by parents' occupation, scored on the Blishen scale, as it was for the Edmonton samples. The means and SDs of scores on the Blishen scale were as follows: 55.85 ± 3.86 (high SES black), 42.84 ± 5.10 (low SES black), 56.44 ± 2.48 (high SES white) and 43.24 ± 5.16 (low SES white). There was a great deal of difficulty in locating appropriate groups of black children, as some school authorities were concerned about the sensitive issue of black/white comparisons. On several occasions the investigators had to bypass the school and contact the children through community organizations. Thus, the sample size for black parents and children who completed all tests could not be increased beyond 20 for each SES; for whites the size was 30 for each SES level.

Tests of cognitive abilities

All children were tested individually on the following tests:

1. Coloured Progressive Matrices. Raven devised this test, which has been widely used with children and is purported to measure reasoning. It is a culture-reduced test, measuring general intelligence.
2. Memory for Designs. Graham and Kendall (1960) devised a memory test for visually presented designs which is supposed to detect minimal brain

damage. It has been used to distinguish normals from retardates (Richie & Butler, 1964). However, in our study it is used mainly as a memory task for designs.

3. Cross-modal coding. The task was modified from one used by Birch and Belmont (1964). Ss listened to patterns of sound, following which they were asked to recognize visually which of the three dot patterns resembled the auditory stimuli. Birch and his associates have found that this test discriminates between good and poor reading ability and between good and poor nutritional status of children. They do not rule out the possibility that the two variables are coexistent (Cravioto, Gaona, & Birch, 1967). The task was specially chosen here to demonstrate S's ability to code auditory materials and recall them visually. It can be classified as a short-term memory task.
4. Visual short-term memory. The visual task consisted of separate presentations of 20 five-digit grids (devised by Drs. Howarth and Brown of the University of Alberta). It was a typical short-term memory task in which a grid was presented for S's viewing for 5 seconds followed by a neutral filler task of color naming for 2 seconds to prevent rehearsal, and then required S to recall the digits on an empty grid. Ss looked at 22 stimulus grids, one after another, and reproduced the digits on an empty grid after each presentation. The first two grids served as practice stimuli.
5. Figure copying. The test was developed by the Gesell Institute (Ilg & Ames, 1964). It has been extensively used and provides an estimate of intellectual ability (Jensen & Rohwer, 1970). It does not involve a memory component, as the child is required merely to copy the geometrical figure while it is in view.
6. Word reading speed. It is obtained from one of the Stroop (1935) charts. The names of four colors (red, green, yellow and blue) are written in block capitals and black ink, each 10 times in a random order. S is timed with the help of a stopwatch for reading the 40 words.

In addition to the above lists, only the Edmonton samples were given an auditory short-term memory test which yielded the serial and free recall scores. The test had 12 four-word lists prepared from a pool of semantically similar words (big, great, fat, high, large, huge, long, tall, wide) and 12 four-word lists of unrelated words (bar, book, cow, day, few, hot, key, pen, wall). All lists were presented by magnetic tape, the semantic and the unrelated lists following each other in a predetermined random order.

Further, the Edmonton sample had scores on reading and mathematics achievement from school records.

Parental attitude and personality tests

The adult personality tests, as well as the interviews, were conducted with the father of the child in an informal manner, mostly at the parents' home. For the children's tests, the same individual administration procedure as was used for cognitive tests was followed. Parents of native children often knew more Cree than English. While interviewing them, the experimenter (J. Manos) was accompanied by a Cree Indian who knew the families on the reserve and was good at both Cree and English. He interpreted the questions for the children after the experimenter presented them in English. The following is a brief description of the various tests used.

1. Index of educational environment (IEE) scale. Parental attitude towards the child's education was measured by using this scale first devised by Wolf and subsequently modified by Dyer (Wolf, 1966; Dyer, 1967). The part of the scale used in the present study elicited parents' responses regarding (i) interest in the child's academic achievement, (ii) knowledge of child's educational progress, (iii) parental aspirations for the education of the child, and (iv) parental preparation for higher education of the child. The parents' responses were obtained on the basis of an interview during which the interviewer asked a fixed set of general questions in each of the four areas. The interviews were taped so that scoring would not have to take place during the interview. Scoring was done on a nine-point scale for each of the four areas of parental attitude and then added to get a total score.
2. Extraversion and neuroticism. Eysenck's (1964) children's scale for measuring extraversion-introversion and neuroticism was administered to the children, while the adults were given the Maudsley Personality Inventory.
3. Rigidity or extreme response set. The scale evolved by Soueif (1958) is an ingenious device to illustrate extreme response set in adults. It is a friendship checklist requiring the subject to consider a series of adjectives, when attributed to a friend, and judge each one on a five-point "absolutely desirable" to "absolutely undesirable" scale. The total number of extreme responses scored as 1 and 5 are then counted for the subject, which is his rigidity score. The children's version of the scale was prepared by changing some of the adjectives so that these could be understood by children. Further research clarifying the nature of this test is to be found in Biggs and Das (1973).
4. Locus of control. The adult measure of locus of control was derived from Rotter (1966). The external-internal locus of control scale (ELC-ILC) measures a person's generalized expectancy regarding the causal nature of "behavior outcome sequences." The children's locus of control scale, devised by Crandall, Kratkovsky and Crandall (1965), was used to assess a child's belief in reinforcement responsibility exclusively in intellectual-academic achievement situations. The intellectual achievement responsibility, or IAR questionnaire as it is called, gives a total internal responsibility score as well as separate subscores for beliefs in internal responsibility for successes and failures. In the present study we have made use of the total internal scores only.

Results and Discussion

Personological characteristics of native parents and low and high SES white parents

For the purpose of comparison among the three parent groups, the means of biographical and personality variables are presented in Table 1. Inspection of Table 1 reveals that the native child's father has, on the average, five years of education and the mother, six years. In this respect they do not differ significantly from low SES white parents.

The native parents, again, have as high an aspiration and interest, etc., for their child's education as the low SES white parents, but both fall short of the aspirations of high SES whites, as might be expected.

It was interesting to note that the native parents were significantly more introverted compared to the two groups of whites; while the reason for this is not

clear, one may assume that some of the items on the questionnaire (e.g., Do other people think of you as being very lovely? Can you easily get some life into a rather dull party?) may not be very meaningful for them. However, their scores on neuroticism and on locus of control were essentially similar to those of white parents. In rigidity or extreme response set the low SES whites were significantly higher than both the native and the high SES whites, who did not

TABLE 1
MEANS AND SDs OF PERSONOLOGICAL AND COGNITIVE VARIABLES:
HIGH AND LOW SES EDMONTON WHITES AND LOW SES CANADIAN
NATIVE GROUP¹

Variable	Group	Mean	SD	Mean Comparisons ²	
Father's education	1	11.18	1.54	1 vs 2	$p < .01$
	2	8.20	2.03	1 vs 3	$p < .01$
	3	5.27	3.07		
	Total	8.22	3.31		
Mother's education	1	10.90	1.38	1 vs 2	$p < .01$
	2	8.57	1.61	1 vs 3	$p < .01$
	3	6.13	2.46		
	Total	8.54	2.68		
IEE interview	1	24.23	5.36	1 vs 2	$p < .01$
	2	15.63	7.50	1 vs 3	$p < .01$
	3	12.40	5.57		
	Total	17.42	7.90		
Adult extraversion	1	11.90	3.40	1 vs 3	$p < .01$
	2	10.93	3.17	2 vs 3	$p < .01$
	3	7.57	1.36		
	Total	10.13	3.32		
Adult rigidity	1	22.23	6.98	1 vs 2	$p < .01$
	2	31.23	11.88	2 vs 3	$p < .01$
	3	18.00	2.27		
	Total	23.82	9.66		
Child extraversion	1	17.87	2.91	1 vs 3	$p < .01$
	2	17.07	3.33	2 vs 3	$p < .05$
	3	15.03	2.50		
	Total	16.66	3.12		
Word reading	1	23.20	4.91	1 vs 3	$p < .01$
	2	24.80	4.37	2 vs 3	$p < .01$
	3	30.47	7.88		
	Total	26.16	6.60		
Progressive Matrices	1	29.37	5.14	1 vs 3	$p < .01$
	2	26.60	4.87	2 vs 3	$p < .05$
	3	23.37	4.70		
	Total	26.44	5.41		
Figure copying	1	16.83	1.97	2 vs 3	$p < .01$
	2	16.17	1.42		
	3	17.57	1.65		
	Total	16.86	1.76		
Cross-modal Coding	1	27.30	3.55	1 vs 3	$p < .01$
	2	26.73	3.52	2 vs 3	$p < .01$
	3	23.27	5.21		
	Total	25.77	4.49		

¹ Groups: 1 = high white (N=60) 2 = low white (N=60) 3 = native (N=60)

² ANOVA and Sheffé's multiple comparisons of means. Nonsignificant differences are not noted.

differ from each other. Their rigidity appears to be the manifestation of their experience of greater tension caused by constant social comparison of themselves with the majority group members. Low SES white parents, contrasted with high SES, may experience greater tension due to the operation of social comparison process in which they keep on striving to attain the levels of high SES white parents. The native parents, however, experience less tension because they do not consider white parents as their comparison group and insulate themselves against external cultural influences that might cause greater tension for them.

*Personality and cognitive processes of the native,
low and high SES white children*

Table 1 presents the means of the three groups of children separately on various tests. Comparison of the means reveals that the only personality difference was in regard to introversion-extraversion. The native children, like their parents, were significantly more introverted than the two groups of white children. Beyond this, the three groups of children had similar personality scores in locus of control, neuroticism and rigidity. Thus, even though the native children were at a reserve school, they were more similar than different on the personality tests used here in comparison to the white children.

Cognitive tests in the battery were divided into three categories in terms of information integration: those that require a simultaneous processing (Progressive Matrices, Figure Copying, Memory for Designs), those requiring successive processing (short-term memory) and those measuring "speed" (word reading speed). The cross-modal coding task has an ambiguous status because it may involve each of the above three processes (Das, 1973b). Mean comparisons of the three groups on cognitive tests as presented in Table 1 reveal that native children were superior to the low SES whites in Figure Copying, obviously reflecting a better spatial-simultaneous process. However, they performed less well in the other simultaneous task, the Progressive Matrices, than white children, irrespective of SES. Their scores on cross-modal coding were also lower. But in successive processing their competence was equivalent to that of the white children. Speed of word reading is very much dependent on facility with English. Since the native children are less proficient in English than the whites, it was not unexpected that in the reading task the whites were faster.

Personological characteristics of black and white parents

The black and white parents belonged to both low and high SES. Thus there were four groups of parents. Is it more likely to find that the low and high SES groups differ than to find that white and black groups differ within the same SES?

An examination of Table 2, which presents the relevant means, indicates that both fathers and mothers of low SES black children, as well as those of low SES white children, had significantly fewer years of schooling than those of the high SES white. The low SES parents, irrespective of ethnicity, had a less favorable educational attitude for their children, particularly when compared with the higher SES white parents. In terms of personality tests, the ethnic or SES groups were indistinguishable except for rigidity. In this the two black groups, irrespective of SES, tended to be more rigid, especially when compared

TABLE 2
MEANS AND SDs OF PERSONOLOGICAL AND COGNITIVE VARIABLES:
HIGH AND LOW SES WHITE AND BLACK MONTREAL GROUPS¹

Variable	Group	Mean	SD	Mean Comparisons ²
Father's education	1	8.50	1.57	1 vs 4 $p < .01$
	2	9.71	1.84	3 vs 4 $p < .01$
	3	8.97	2.06	
	4	10.65	1.01	
	Total	9.53	1.83	
Mother's education	1	7.75	1.59	1 vs 4 $p < .01$
	2	9.01	1.75	3 vs 4 $p < .05$
	3	8.97	2.36	
	4	10.31	1.29	
	Total	9.14	1.99	
IEE interview	1	18.10	4.52	1 vs 4 $p < .05$
	2	19.35	5.28	3 vs 4 $p < .01$
	3	18.03	5.81	
	4	22.27	4.26	
	Total	19.58	5.25	
Adult rigidity	1	25.00	7.09	1 vs 4 $p < .01$
	2	24.25	12.71	2 vs 4 $p < .01$
	3	20.47	10.59	
	4	14.53	9.25	
	Total	20.35	10.73	
Child IAR	1	24.20	4.19	1 vs 4 $p < .05$
	2	24.40	4.15	
	3	25.20	3.75	
	4	27.33	3.56	
	Total	25.48	3.99	
Figure copying	1	9.70	3.16	1 vs 4 $p < .01$
	2	11.20	1.54	3 vs 4 $p < .02$
	3	10.73	2.50	
	4	12.90	2.86	
	Total	11.27	2.81	
Visual short-term memory	1	86.15	9.71	1 vs 4 $p < .05$
	2	89.90	10.31	
	3	90.10	10.02	
	4	94.00	6.96	
	Total	90.44	9.43	

¹ Groups: 1 = low black (N=20) 3 = low white (N=30)
2 = high black (N=20) 4 = high white (N=30)

² ANOVA and Scheffé's multiple comparisons of means. Nonsignificant differences are not noted.

with the high SES whites. Thus, Soueif's (1958) original finding regarding the higher rigidity scores for minority groups under tension appears to be supported in this case.

Personality and cognitive processes of black and white children

As is revealed in Table 2, with respect to personality patterns, the four groups of children were indistinguishable from each other except in one test; this was the locus of control (IAR). The high SES white children were significantly more internally controlled than the low SES black children. The remaining groups were not significantly different.

In cognitive tests reflecting simultaneous processing, there was no difference between the four groups except in Figure Copying. Low SES blacks and whites had significantly lower scores on this test than high SES whites.

The Progressive Matrices, particularly, would be expected to differentiate the groups on the basis of SES (Jensen, 1970). It did not do so either in the Edmonton or in the Montreal samples.

In short-term memory, a test of successive processing, the low SES black children had a lower score than the high SES whites; however, the difference was barely significant. Performance on the speed test was similar for the four groups.

On the whole, it appears that the native children can be more readily distinguished from the whites in terms of personality traits and in cognitive processes than either the black and white samples, or the high and low SES samples.

Prediction of Progressive Matrices performance of white and black children

The Progressive Matrices was the only cross-cultural intelligence test that we used across the subcultural groups. To what extent do noncognitive factors correlate with the performance on this test? Parental attitude, Blishen SES scores of parents *within each SES group* and the locus of control (IAR) of the child were employed as predictor variables for this purpose.

The percentage of variance accounted for by parental attitude, SES and locus control as a result of stepwise regression analysis is shown in Table 3. First, consider the Edmonton sample—only the high SES white children's scores are predictable. The major predictor is parents' attitude. Parents' SES standing adds to the predictability (r between parent's attitude and SES is .25, which is low enough, so that one does not mask the effect of the other). The contribution of the child's locus of control does not improve predictions substantially. For the high SES black sample in Montreal, parents' attitude is the best

TABLE 3
PREDICTION OF PROGRESSIVE MATRICES SCORES BY PARENTAL
ATTITUDE (PA), SES AND CHILD'S LOCUS OF CONTROL (LC)

Group			Percentage of Variance			
<u>Edmonton</u>						
Low SES white	4.95	PA	18.21	PA + SES	21.77	PA + SES + LC
High SES white	21.53**	PA	27.21**	PA + SES	30.22*	PA + SES + LC
Native	2.59	PA	---		4.48	PA + LC
<u>Montreal</u>						
Low SES white	5.12	PA	6.31	PA + LC	6.34	PA + LC + SES
High SES white	23.13**	LC	25.66*	LC + SES	25.85*	LC + SES + PA
Low SES black	0.90	SES	1.06	SES + LC	1.07	SES + LC + PA
High SES black	39.36**	PA	42.61**	PA + LC	43.64*	PA + LC + SES

* $p < .05$

** $p < .01$

predictor of Progressive Matrices. The next best predictor is child's locus of control. SES adds very little to the total variance; however, this should be interpreted cautiously because of the moderate correlation of .59 between locus of control and SES. For the high SES white sample, child's locus of control is the best predictor, followed by SES; parents' attitude adds almost nothing to the variance.

Two points emerge from these results: (1) the Progressive Matrices scores of none of the low SES groups—whites and blacks—could be predicted; (2) in regard to the high SES groups, reasoning ability as measured by the Matrices is predictable from parental attitude and locus of control variables.

Prediction of school achievement

Reading and mathematics achievement test scores (school achievement) were available for the white children in Edmonton. The multiple regression analyses with respect to prediction of school achievement are presented in Tables 4 and 5. Multiple R was computed, the predictor variables being parents' attitude, SES scores and child's locus of control. It was found that reading achievement could not be predicted for the high SES children and mathematics achievement was only marginally predictable by parents' attitude. However, for low SES children both achievements were highly predictable. Parents' attitude was the most important predictor, accounting for almost 40 percent of the variance in mathematics achievement and 20 percent in reading achievement. The parents' standing on the SES scale added substantially to these predictions, but the child's locus of control scores did not (*r* between parent's attitude and SES was .55 based on 30 cases; *r* between locus of control and parental attitude or locus of control and SES were near zero). Thus, the proficiency of the low SES child in reading and mathematics is related to both parental attitude and SES. But his Progressive Matrices performance was not predictable. For high SES white children the opposite seemed to be true. These observations should be somewhat tempered by the fact that the standard deviation for reading achievement was larger for the low SES children than for the high SES (18.18 and 12.25 respectively). But they were comparable for mathematics achievement (16.86 and 15.88).

The conclusion appears to be that school achievement in the lower class children is substantially related to parental attitude and the position of their

TABLE 4
PREDICTION OF READING ACHIEVEMENT AND MATHEMATICS
ACHIEVEMENT BY PARENTAL ATTITUDE (PA), SES AND CHILD'S LOCUS
OF CONTROL (LC)

Group	Percentage of Variance					
	Reading			Mathematics		
High SES white	10.81	12.00	12.42	16.18*	17.55	18.28
	PA +	LC +	SES	PA +	LC +	SES
Low SES white	20.24**	35.76**	35.92**	39.64**	45.54**	45.82**
	PA +	SES +	LC	PA +	SES +	LC

* *p* < .05

** *p* < .01

parents on the SES scale, even within the narrow range of low SES scale values. A more specific study should inquire into the pattern of interaction between parents and their children of elementary school age in low SES homes, with particular attention to school work, in order to understand the basis for these relationships.

TABLE 5
PREDICTION OF READING ACHIEVEMENT AND MATHEMATICS
ACHIEVEMENT SCORES BY COGNITIVE TESTS¹

Group	Percentage of Variance							
High SES white	21.62**	27.54*			16.98*	22.85*		
	FC	+	STM		FC	+	CMC	
Low SES white	24.16**	29.85**	33.53**	35.69*	20.73**	24.64*	26.58*	
	SR	+	STM	+	FC	+	CMC	

¹Figure Copying, STM Visual, Cross-Modal Coding, Serial Recall and Free Recall

* $p < .05$
** $p < .01$

In Table 5, the predictor variables are serial and free recall, short-term memory, Figure Copying and cross-modal coding. Of these, the memory tests require successive processing, Figure Copying uses simultaneous processing, and Cross-modal Coding uses both. The Cross-modal Coding task had been found to correlate highly with reading ability in previous research (Birch & Belmont, 1964; Birch & Gussow, 1970). One would expect that successive processing will be closer to reading, whereas simultaneous will better predict mathematics achievement. The results did not quite confirm the expectation. For the high SES children the free recall and short-term memory scores were good predictors of reading, as expected. The predictor variables for low SES reading achievement were serial recall and short-term memory, and to a lesser extent Figure Copying and Cross-modal Coding. In other words, a general competence in the cognitive tests predicted reading. Mathematics achievement was marginally predicted by Figure Copying and Cross-modal Coding for the high SES group. The prediction was somewhat better for the low SES group, jointly shared by serial recall, Figure Copying and Crossmodal Coding. It was interesting to note that serial recall was the major predictor of mathematics achievement in the low SES group. Perhaps this suggests that the low SES children are depending on a successive strategy in doing mathematics, whereas the high SES children adopt the expected simultaneous strategy. In predicting reading, Cross-modal Coding was not an important variable; this result should force one to rethink the association between the two variables.

It appears, then, that the social-personality variables predicted the mathematics achievements of low SES children much more efficiently than the cognitive tests, and both kinds of variables predicted reading achievement equally well.

In conclusion, we wish to point out the role of SES and parental attitude as predictors of reasoning and school achievement. Past evidence (Wolf, 1966) has

demonstrated that parental attitude towards the child's education correlated to an extent of 0.69 with child's IQ; in contrast, the correlation with SES was between 0.20 and 0.40. Our results generally showed parental attitude to be a better predictor than SES.

School achievement is another variable that has been associated with parental attitude and SES (Wolf, 1966; Dyer, 1967). Again, parental attitude has been found to correlate higher with school achievement than SES. Our data are consistent with those of Wolf and Dyer in this respect, although we did not use their full scale. The same general consensus is expressed in the report of the Plowden Committee (1967)—they found that the variation in the children's school achievement and IQ is specifically accounted for more by the variation in parental attitudes than by either the variation in the material circumstances of parents or by the variation in schools.

This research was supported by a Canada Council Grant (S69-0749) to J. P. Das. Dr. R. Kanungo was the other investigator and Dr. J. Manos, then a graduate student at the University of Alberta, was the senior research assistant for the project.

References

- Biggs, J., & Das, J. P. Extreme responses set and modes of maximizing cognitive clarity. *British Journal of Social and Clinical Psychology*, 1973, 12, 199-210.
- Birch, H. G., & Belmont, L. Auditory-visual integration in normal and retarded readers. *American Journal of Orthopsychiatry*, 1964, 36, 852-861.
- Birch, H. G., & Gussow, J. D. Disadvantaged children. New York: Grune & Stratton, 1970.
- Blishen, B. R., Jones, F. E., Naegele, K. D., & Porter, J. *Canadian Society*. Toronto: Macmillan, 1965.
- Cole, M., & Bruner, J. S. Cultural differences and inferences about psychological processes. *American Psychologist*, 1971, 26, 867-876.
- Crandall, V. C., Kratkovsky, W., & Crandall, V. J. Children's beliefs in their own control of reinforcements in intellectual academic achievement situations. *Child Development*, 1965, 36, 91-109.
- Cravioto, J., Gaona, C. E., & Birch, H. G. Early malnutrition and auditory-visual integration in school age children. *Journal of Special Education*, 1967, 2, 75-82.
- Das, J. P. Cultural deprivation and cognitive competence. In N. R. Ellis (Ed.), *International review of research in mental retardation*. Vol. 6. New York: Academic Press, 1973. (a)
- Das, J. P. Structure of cognitive abilities: Evidence for simultaneous and successive processing. *Journal of Educational Psychology*, 1973, 65, 103-108. (b)
- Deutsch, M., and associates. *The disadvantaged child*. New York: Basic Books, 1967.
- Dyer P. B. The effects of environmental variables on the achievement of elementary school children in Trinidad. Unpublished Ph.D. dissertation, University of Alberta, 1967.
- Eysenck, S. B. G. A new scale of personality measurement in children. *British Journal of Educational Psychology*, 1964, 34, 197-201.
- Graham, F. K., & Kendall, B. S. Memory-for-Designs Test: Revised general manual. *Perceptual and Motor Skills*, Monograph Supplement 2-7, 1960, 11, 147-188.
- Ilg, F. L., & Ames, L. *School readiness*. New York: Harper, 1964.
- Jensen, A. R. Reducing heredity-environment uncertainty: A reply. *Harvard Educational Review*, 1969, 39, 449-482.
- Jensen, A. R. A theory of primary and secondary familial mental retardation. In N. R. Ellis (Ed.), *International review of research in mental retardation*. Vol. 4, New York: Academic Press, 1970, 33-100.

- Jensen, A. R. A two-factor theory of familial mental retardation. *Human Genetics*. Proceedings of the 4th International Congress of Human Genetics, Paris, September 1971.
- Jensen, A. R., & Rohwer, W. D. An experimental analysis of learning abilities in culturally disadvantaged children. Final Report, 1970, OEO Project No. 2404, U.S. Office of Economic Opportunity.
- Luria, A. R. Towards the problem of the historical nature of psychological processes. *International Journal of Psychology*, 1971, 6, 259-272.
- Plowden, B. *Children and their primary schools*. A report of the central advisory council for education (England). London: Department of Education and Science, 1967.
- Richie, J., & Butler, A. Performance of retardates on the Memory-for-Designs test. *Journal of Clinical Psychology*, 1964, 20, 108-110.
- Rohwer, W. D., Jr. Learning, race, and school success. *Review of Educational Research*, 1971, 41, 191-220.
- Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs*, 1966, 80, 1.
- Sherif, M. Group influences on the formation of norms and attitudes. In T. M. Newcomb & E. L. Hartley (Eds.), *Readings in social psychology*. New York: Holt, Rinehart & Winston, 1947.
- Soueif, M. I. Extreme response sets as a measure of intolerance of ambiguity. *British Journal of Psychology*, 1958, 49, 329-334.
- Stroop, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 1935, 18, 643-661.
- Wolf, R. The measurement of environment. In A. Anastasi (Ed.), *Testing problems in perspective*. Washington, D.C.: American Council on Education, 1966, 491-503.

MARY NIXON

and

L. R. GUE

The University of Alberta

Women Administrators and Women Teachers: A Comparative Study

This exploratory study was designed to compare three groups within the teaching profession—women school administrators, women teachers matched with administrators on years of teaching experience and teacher education, and women teachers randomly selected—on variables which may be relevant to career decisions. Literature on professionalism and professional role orientation, on socialization and sex role orientation, provided the conceptual framework for the study. (Dr. Nixon is a sessional instructor and Dr. Gue is Associate Professor in the Department of Educational Administration at The University of Alberta.)

The under-representation of women in positions of administrative responsibility in proportion to their numbers in the teaching force is a phenomenon which is not confined to Alberta (Asper, 1974, p. 5; McIntosh, 1974, p. 28). Certainly, in Alberta, the situation appears to be getting worse rather than better. The percentage of women teachers who are principals is declining while the percentage of men in administrative and supervisory positions is increasing (Alberta Teachers' Association, 1973, p. 2). In Alberta, in general, the more administrative responsibility attached to a position the less the probability that the position will be held by a woman (Nixon & Hrynyk, 1973, p. 1). This decrease in the number of women in school administrative positions is taking place at a time when there is concern that the 167 recommendations of the Royal Commission on the Status of Women (1970) intended to eliminate discrimination against women in all areas of Canadian life are being somewhat slowly implemented.

It is of some importance to know whether teachers perceive attitudinal barriers which deter them from applying for school administrative positions or whether, as Alberta school superintendents believe, lack of applications from women teachers is the result of "lack of interest" in the administrative task (Nixon & Hrynyk, 1973, p. 25).

Ensuing from present practice is a waste of potential leadership talent. At present educational leadership is drawn from approximately fifty percent of the teaching force. As Dale (1973) points out:

At a time when good and perhaps great leadership is needed in education, the pool of leadership development must be expanded to include women or we must all learn to live with the mediocrity which is certain to result from a restricted source of leadership talent (p. 125).

There would appear to be a relationship between professionalism and upward mobility. High scores on a professional role orientation scale have been shown to be associated with persons occupying the higher positions in the educational hierarchy (Hrynyk, 1966, p. 210). The Professional Role Orientation scale devised by Hrynyk (1966) has shown that in Alberta there were differences in the professional orientations of different groups of teachers. Male teachers scored higher than did women teachers (p. 174) and "in general, higher scores were associated with positions which are usually ascribed higher 'status' in the school hierarchy" (p. 210).

An explanation has been sought in the socialization and resulting sex-role orientations of women for women teachers' lower professional role orientation scores and their seeming lack of interest in "status" positions. From very early childhood parents are concerned that their children learn appropriate sex role behavior (Goldberg & Lewis, 1972, p. 33). Most women acquire in childhood a traditional definition of the female role as that of wife and mother. Stereotyped sex roles are reinforced by formal schooling (Royal Commission on the Status of Women, 1970, pp. 180-183; McLure, 1973, pp. 111-113; Shack, 1973, p. 62). Epstein (1970) believes that childhood socialization "is nearly always crucial to later occupational decisions" (p. 19). Encouraged to take courses which guide them into traditionally female occupations, girls who hope to combine professional and homemaker roles may perceive teaching as a realistic choice.

It has been suggested that for some women achievement in the world of work is something which is better avoided since it is threatening to their sense of femininity (Horner, 1972, p. 173; Kimball, 1973, p. 132). For such women "success" in the man's world of work has a negative value. Women are warned, whether overtly or not is unimportant, not to be too successful in the man's world of work. However, if the career woman is often considered aggressive, overly ambitious, and by implication unfeminine, the homemaker is often held in low esteem. Society pays lip service to the role of woman as homemaker (Kimball, 1973, p. 132) but apparently very little more. Among women, it is increasingly recognized that the wife who is unable to define herself in terms of an occupational role in the labour force may find herself "ciphered out as an individual" (Bernard, 1971, p. 154). The response of many women to such ambivalent attitudes may well be a dissatisfaction with the role of homemaker together with a latent fear of appearing too successful in a career outside the home. The female teacher who believes her role of homemaker to be of greater importance than her professional role is unlikely to have any particular ambition to further her career by either seeking or wanting an administrative appointment.

Another explanation of women teachers' reluctance to apply for administrative positions (Hrynyk & Nixon, 1973) may be the belief that applications from women for administrative appointments are unlikely to meet with

success. Fishel and Potter (1973) would argue that "women teachers have seen how difficult it is for a woman to obtain an administrative appointment and have stopped trying to do so" (p. 390).

Research Procedures

This study was carried out in Edmonton, Alberta, Canada, during March, April and May, 1974. Permission was sought to approach women teachers and administrators in the Edmonton Public Schools for their cooperation in the study. The Edmonton Public School Board approved the study and provided information on all women teachers deployed in the system.

The Sample

Originally the sample consisted of one hundred and fifty-nine administrators and teachers in the Edmonton Public School System: (1) the total population of 53 women principals and assistant principals, (2) an equivalent number of women teachers matched with the administrators by a pairing system on factors of years of teacher education and teaching experience, and (3) a third numerically equivalent group drawn, using a table of random numbers, from a computer printout of all women teachers in the Edmonton Public School System. By holding constant two variables associated with administrative positions—teacher education and teaching experience—the inclusion of the matched group made possible a more careful examination of other variables on which administrators and teacher groups might differ.

The initial number of administrators and teachers were contacted by a letter dated March 1, 1974, which explained the purpose of the research. Out of a total of 159 administrators and teachers who were asked to participate in the study, 147 (92.4 percent) agreed to do so. However, actual participation rates were lower; 133 administrators and teachers completed and returned the questionnaire. Of these 133 participants, 47 were school administrators, 44 were in the group matched with administrators on years of teacher education and teaching experience, and 42 were in the randomly selected group.

Composition of the Interview Sample. Ten participants from each group (five who obtained a "traditional" score on the Sex Role Orientation scale and five who obtained a "liberationist" score on the Sex Role Orientation scale) were asked to take part in an interview situation. All agreed to do so.

Data Collection

All those who agreed to take part in the study were assigned a code number. During the third week in March a three-part questionnaire was mailed to all participants together with a stamped, self-addressed envelope. Part A sought responses on: (1) biographical data, (2) whether administrative positions had been or would be applied for, (3) personal and school board preferences for men as school administrators, and (4) time as a factor in accepting or rejecting an administrative appointment. Part B of the questionnaire consisted of a Professional Role Orientation scale (Hrynyk, 1966) and Schmidt's (1973) Sex Role Orientation scale made up Part C. In addition, free space was left on the questionnaire for comments.

The first section of the questionnaire consisted of closed, mutually-exclusive categories. Participants were asked to check their responses to the information sought. For the questions on: (1) personal and school board

preferences for male administrators, (2) time as a factor in accepting or rejecting an administrative appointment, and (3) interest in making an application for an administrative position, three response categories were provided. These consisted of Yes, No, and Undecided.

The Professional Role Orientation scale was made up of Likert-type items with response categories, Agree Strongly, Agree, Undecided, Disagree and Disagree Strongly. The response categories were weighted 1-5, the direction of the weighting being dependent upon the item.

The Sex Role Orientation scale was also made up of Likert-type items. The response categories, Strongly Agree, Somewhat Agree, Undecided, Somewhat Disagree, Strongly Disagree were weighted 1-5, the direction of the weighting being dependent on the item.

A check-list and a tape-recorder were both used in the interview situation. Questions were open-ended and the responses were categorized later from the tape recordings made at the interviews. Information was sought on: (1) socialization and career choice, (2) social change and sex roles, (3) women in administration, (4) stereotyping, and (5) application/non-application of participants for administrative positions.

Validity. Trust is crucial to the validity of the responses in this type of research where respondents are being asked to give information about their personal attitudes and beliefs. Essentially this is a problem of establishing trust on the part of the participants that the information which they volunteer will not be used to their detriment and that confidentiality of the responses will be safeguarded. Glazer (1972, p. 11) contends that this problem of acceptance is common to all field research.

In an attempt to gain the acceptance of which Glazer writes, the following steps were taken:

1. A letter outlining the purpose of the study and the method of data collection was sent to each member of the administrative, matched and random groups.
2. Each participant was a volunteer.
3. Each participant was assured of anonymity.
4. An attempt was made to ensure that the content of the items was personally non-threatening in the hope that this would lower the tendency for participants to give socially desirable responses.

The Professional Role Orientation (PRO) scale and the Sex Role Orientation (SRO) scale were accepted as valid instruments. The PRO scale developed by Hrynyk (1966) has been further validated by Scharf (1967) and Veitch (1969). The SRO scale has a pragmatic validity in that Schmidt claims that the scale does differentiate between traditionalists and liberationist attitudes (1973, p. 78).

Treatment of the Data

Questionnaire responses and categorized interview data were translated to computer cards. No attempt was made to compensate for missing responses in Part A of the questionnaire or in the interview data but missing responses on the PRO and SRO scales were assigned a value of 3 (undecided). A measure of professionalism was obtained by computing the sum of all items. Scores on the subscales provided the ratings for five dimensions of professionalism. Parti-

cipants with a composite score which fell in the upper third of all PRO scores were defined as "high" professionally oriented teachers and all those with scores in the lower third of all PRO scores were defined as "low" professionally oriented teachers. Similar procedures were followed for the SRO scale. Participants with scores which fell in the upper third of all scores were defined as having a "traditional" sex role orientation; those whose scores fell in the lower third of all scores were defined as having "liberationist" sex role orientations.

The focus of the study was a comparison between administrators, matched and random teacher groups and throughout the statistical analysis the three groups have been treated as three samples drawn from the populations. A one way analysis of variance was one statistical procedure employed in order to test for significant statistical differences between populations. The *F* test was used for intergroup comparisons. Where *F* proved to be significant beyond the .05 level, the nature of the implied difference was investigated further by the application of the Scheffé method of posteriori comparisons. For intragroup comparisons the T test of independent samples was used. Where the level of measurement of the dependent variable did not allow the use of parametric statistics, the statistic chi-square was used as a test for significance of difference.

Results

Questionnaire Data

Biographical variables. With respect to the biographical variables of teacher education, teaching experience and the related variable of age, there was a statistically significant difference between the random teacher group and the administrative and matched groups. Teachers in the random group had less years of teacher education, less years of teaching experience and were younger than participants in the other two sample groups.

A chi-square test of significance revealed significant difference between the sample groups in marital status. The distribution of the dichotomous variable married/not married is presented in Table 1. Although more than 70 percent of the teachers in the random group were married, this was true of less than 50 percent in the administrative group. This finding was supportive of the research reported by Epstein (1979, p. 96) that the incidence of unmarried women was greater among those "who rose to the top of their profession." For women who are not married, professional responsibilities do not have to com-

TABLE 1
CHI-SQUARE TEST OF SIGNIFICANCE BETWEEN ADMINISTRATORS,
MATCHED AND RANDOM TEACHER GROUPS CLASSIFIED
ACCORDING TO MARITAL STATUS

Marital Status	Administrators		Matched		Random	
	N	%	N	%	N	%
Married	23	48.9	26	59.1	31	73.8
Not Married	24	51.1	18	40.9	11	26.2
TOTAL	47	100.0	44	100.0	42	100.0

$\chi^2 = 5.756$ (significant at the .05 level) d.f. = 2

pete with other role obligations of “wife” and “mother”—roles which are traditionally regarded as full-time occupations. Despite the difficulties presented by competing roles commented upon by some women teachers, no relationship between marital status and unwillingness to apply for administrative positions was established. Neither marriage nor dependents of any age precluded women teachers from applying or intending to apply for administrative positions.

Professional Role Orientation. Table 2 contains the results of the analysis of variance of the Professional Role Orientation scores between groups and the Scheffé multiple comparison of means. The *F* value of 4.70 with a probability of .01 indicated that the differences between the means was statistically significant. The Scheffé multiple comparison of means revealed a significant difference in group mean scores between the administrators and the random teacher group.

TABLE 2
ONE-WAY ANALYSIS OF VARIANCE OF PROFESSIONAL ROLE ORIENTATION
SCORES AMONG ADMINISTRATORS, MATCHED AND RANDOM
TEACHER GROUPS FOLLOWED BY THE SCHEFFE MULTIPLE COMPARISON
OF MEANS

Source	s.s.	m.s.	d.f.	<i>F</i>	<i>p</i>
Between Groups	0.873	436.5	2	4.70	.01
Within Groups	0.120	92.83	130		

Scheffé Multiple Comparison of Means			
Groups	Administrators	Matched	Random
Means	102.0	100.0	95.83
Administrators	-		
Matched	0.64	-	
Random	0.01*	0.126	-

* Significant beyond the .05 level.

A willingness to spend additional time was presumed to be related to the service attribute of professionalism. Responses to the question, “Would you turn down an administrative position if you thought it was going to be more time-consuming than classroom teaching?” are shown in Table 3. Discrepancy in the *N*s between Table 3 and other tables is due to the arbitrary omission of “undecided” responses. A chi-square test revealed significant differences between the groups in their willingness to spend, if necessary, additional time on the administrative task. More than 90 percent in the administrative group are willing to spend additional time compared to approximately 70 percent in the other two sample groups.

Statistically significant differences existed on the PRO scale between administrators and the random teacher group. This provided additional support for Hrynyk’s finding of a positive relationship between higher status positions

TABLE 3
CHI-SQUARE TEST OF SIGNIFICANCE BETWEEN GROUPS ON RESPONSES
TO QUESTION 15 ON THE QUESTIONNAIRE¹

	Administrators		Matched		Random	
	N	%	N	%	N	%
Yes	3	7.5	12	30.8	11	28.9
No	37	92.5	27	69.2	27	71.1
TOTAL	40	100.0	39	100.0	38	100.0

$\chi^2 = 7.659$ (significant at the .05 level) d.f. = 2

¹ Question 15: Would you turn down an administrative position if you thought it was going to be more time-consuming than classroom teaching?

in the educational hierarchy and higher PRO scores (1966, p. 210). However, the administrators in this study also differed significantly from the random teacher group with respect to years of teacher education. The intervening variable between status positions and PRO scores may well be teacher education. Although the administrators' mean score on the PRO scale was higher than the mean score of the control group for teachers matched with administrators on years of teacher education and teaching experience, this difference in mean scores was not statistically significant. Simpson and Simpson (1969) have remarked that strong enthusiasm is not typical of women teachers' attitudes toward the profession but they do believe that "professional status and extensive training . . . increase women's vocational commitment" (p. 216). Thus professional education rather than the status positions per se may account for the statistically significant differences between the administrators and the random teacher group on the PRO scale.

Within the matched and random sample groups a relationship between willingness to apply for administrative positions and high PRO scores was not established but there was a tendency for the mean scores of those willing to apply for administrative appointments to be higher than the mean scores of those who were not willing to apply for administrative appointments.

Sex Role Orientation. In the development of the theory underlying the testing for statistical differences between groups on the Sex Role Orientation scale was the contention that women who held or who sought line positions would differ in their sex role orientation from those women who neither held them nor sought them. In addition, it was hypothesized that high PRO scores would be associated with less traditional definitions of the female role (i.e., low SRO scores).

No statistical differences were revealed between the sample groups but the random teacher SRO mean was the lowest of the three sample groups. This was in keeping with Schmidt's (1973, p. 106) finding that age was strongly related to sex role attitudes, the younger age categories being more liberationist in orientation.

Within the random teacher group, those who were willing to apply for administrative positions had a statistically significantly lower SRO score denoting a more liberationist sex role orientation, than those participants who were not willing to seek such positions (Table 4). Thus the premise that women

who held or who sought line positions would differ in their sex role orientation from women who had neither held them nor sought them was established within the random teacher sample.

TABLE 4
ONE-WAY ANALYSIS OF VARIANCE OF SEX ROLE ORIENTATION SCORES
WITHIN THE MATCHED GROUP AND RANDOM TEACHER SAMPLES
CLASSIFIED ON THE BASIS OF "APPLY" (GROUP 1) AND "NOT APPLY"
(GROUP 2) FOR ADMINISTRATIVE POSITIONS

Sample Group	Group 1		Group 2		Group 1 S.D.	Group 2 S.D.	d.f.	T ¹	T'
	N	\bar{x}	N	\bar{x}					
Matched	14	146.86	30	146.93	19.39	14.79	42	-0.01	
Random	10	126.30	32	144.34	7.35	19.03	40		-4.41**

¹Where the F test differences between variances proved to be significant ($\geq .05$) the Welch T' has been reported.

Beliefs Regarding Administration. Analysis of the data revealed no statistically significant differences between the groups as to beliefs concerning the preferences of school boards for male administrators. Women administrators appeared to be a little more convinced than other groups of school boards' open and unbiased policies toward them. Just under 20 percent of the administrators believed that men were preferred by school boards and this percentage dropped to 11 percent for the random teacher group. Asper, (1974, p. 88) in a study of women teachers and administrators in Manitoba, reported similar findings. In response to the item, "A school board will encourage male teachers rather than female teachers to apply for administrative positions," 68 percent of teachers either agreed or strongly agreed with the statement. This was also true of 55 percent of women administrators.

Within the matched teacher group a statistically significant relationship was established between SRO scores and preferences for a male authority figures (Table 5). In this group forty-three percent of those with traditionalist SRO scores preferred men in supervisory positions whereas all those with liberationist SRO scores expressed no preference.

TABLE 5
CHI-SQUARE TEST OF SIGNIFICANCE BETWEEN HIGH SRO SCORES AND
LOW SRO SCORES AND PREFERENCE FOR MEN IN SUPERVISORY
POSITIONS WITHIN THE MATCHED¹ TEACHER GROUP

	High SRO		Low SRO	
	N	%	N	%
Prefer men in supervisory positions	7	43.8	0	0.0
No preference	9	56.3	10	100.0
TOTAL	16	100.0	10	100.0

χ^2 corrected for continuity = 3.97 (significant beyond the .05 level); d.f. = 1

¹"Undecided" responses have been omitted in the calculation of this table.

Interview Data

There was no apparent difference in the socialization of any of the sample groups. For the majority of participants in all groups, teaching had not represented so much a "calling" as a practical career choice. In all three sample groups a number of participants said they would consider leaving teaching were they given the opportunity to do so.

Only in the administrative group would a majority of the participants recommend teaching as a career for their daughters. The reasons for administrators' recommendations varied from the pragmatic to the ideal of service as the following quotations illustrate: from "I think it is a marvelous thing for a young person to have that teaching certificate, especially a young lady," to "You can help children and people and society . . ." Generally, however, the reasons given were related to administrators' own personal sense of satisfaction with their career choice.

Regardless of the sample group to which they belonged, few participants believed that teaching provided equal opportunities for men and women to obtain administrative appointments. Beliefs in restricted professional opportunities for women were largely based on the number of women in the upper levels of the educational hierarchy. Administrators differed significantly from the random teacher group in that administrators perceived the position of supervisor as being open to them.

There were no widely held stereotypes of male and female school principals. This might be, in part at least, because few participants had worked in a school where the principal was female and therefore had no experiential base on which to generalize.

Administrators differed significantly from the matched teacher group in their perception of the administrative task. While administrators saw it as an interesting, professional challenge, the manner in which the teacher groups were apt to define the administrative task explained why a number of teachers—especially those in the matched teacher group—had no interest in it.

Despite the commonly made assertion that the homemaker role prevents women teachers from applying for administrative positions, only one participant in each of the teacher groups cited family responsibilities as the most important factor in making the decision not to apply for such a position.

Career decisions of administrators were very much influenced by authority figures in the educational hierarchy. Without some kind of professional sponsorship many expressed the doubt that they would ever have made the decision to apply for administrative positions. It would, therefore, be difficult to overestimate the importance of this variable in helping women teachers to make career decisions.

Conclusions

To generalize widely from an exploratory study such as this would be unwise. The conclusions that follow are both place bound and time bound in that the findings relate to the Edmonton Public School District No. 7 in 1974.

This study gave little support to the inference that women teachers perceive their additional roles of "wife" and "mother" as preventing them from applying for administrative positions. There were no significant differences between those who were willing to apply for administrative positions with respect to

these variables. However, some participants did indicate that their "willingness to apply" was future oriented—to a time when their family responsibilities would be less.

Administrators scored significantly higher on the Professional Role Orientation scale (PRO) than did teachers randomly selected but they did not obtain significantly higher scores than those participants in the matched teacher group where years of teacher education and teaching experience were controlled. This would seem to suggest that years of teacher education and teaching experience might well be important intervening variables between administrative positions and higher scores on the PRO scale.

In general, administrators appeared to obtain more satisfaction from their professional role than participants in the teacher groups. In comparison with teachers, administrators were: (1) less likely than participants in the teacher groups to consider an alternative career to teaching, and (2) more likely than participants in teacher groups to be positive in their recommendation of teaching as a career to others. While a career in school administration was considered as an interesting, professional challenge by administrators, this view was not widely shared among participants in the other sample groups. Teachers' definitions of the administrative task may contribute to their reluctance to apply for such positions.

The majority of participants in all sample groups believed opportunities for women teachers to be restricted. Formal equality within the profession with respect to salaries and fringe benefits has done little to alter the perceptions of the majority of women teachers that opportunities for advancement within the profession are not equal for both sexes. Women teachers hold few expectations for advancement in their careers and many have withdrawn from what they perceive as highly competitive achievement situations. Without some encouragement it is probable that the number of women teachers who apply for administrative positions will continue to be small in proportion to the number of women in the teaching force.

The findings of the study imply the need for imaginative leadership programs by school boards, professional associations and faculties of education, aimed at women teachers. Such programs would serve the purpose not only of identifying potential administrative talent but also of arousing interest in the administrative task as something more than just "paper work, discipline, timetabling, and things like that." In addition, sponsorship of women teachers who are interested in taking professional training in administration would add to a school board's credibility that the appointment of women to positions of administrative responsibility is more than mere tokenism. Entrenched attitudes are slow to change. It will take time, demonstrated fair promotion policies and patience on the part of school boards before women teachers become convinced that they too may attain influential positions in the educational hierarchy.

References

- Alberta Teachers' Association. *The Alberta teaching force*, September, 1972. Supplement to ATA Research Monograph #21. Edmonton: Alberta Teachers' Association, 1973.
- Asper, L. B. Factors affecting the entry of women into administrative positions of the Manitoba school system. Unpublished M.Ed. thesis, University of Manitoba, 1974.
- Bernard, J. The paradox of a happy marriage. In V. Gornick & B. K. Moran (eds.), *Women in sexist society: Studies in power and powerlessness*. New York: Basic Books, 1971, 145-162.
- Dale, C. T. Women are still missing persons in administrative and supervisory jobs. *Educational Leadership*, 1973, 31 (2), 123-127.
- Epstein, C. F. *Woman's place. Options and limits in professional careers*. Berkeley, California: University of California Press Limited, 1970.
- Fishel, A., & Potter, J. Women lose out. Is there sex discrimination in school administration? *The Clearing House*, 1973, 47, March, 387-391.
- Glazer, M. *The research adventure*. New York: Random House, Inc., 1972.
- Goldberg, S., & Lewis, M. Play behaviour in the year-old infant. In J. M. Bardwick (Ed.), *Readings in the psychology of women*. New York: Harper and Row, 1972.
- Horner, M. S. Towards an understanding of achievement-related conflicts in women. *Journal of Social Issues*, 1972, 28, 157-175.
- Hrynyk, N. P. Correlates of professional role orientation in teaching. Unpublished Ph.D. thesis, University of Alberta, 1966.
- Kimball, M. M. Women and success: A basic conflict. In M. Stepehson, (Ed.), *Women in Canada*. Toronto: New Press, 1973, 119-135.
- McIntosh, J. C. Differences between women teachers who do and do not seek promotion. *The Journal of Educational Administration*, 1974, 12 (2), 28-41.
- McLure, G. T. Eliminate sex bias in the curriculum. *Educational Leadership*, 1973, 31 (2), November, 111-113.
- Nixon, M. Women administrators and women teachers: A comparative study. Unpublished Ph.D. thesis, University of Alberta, 1975.
- Nixon, M., & Hrynyk, N. P. Women in school administration. A study of the deployment of women in positions of administrative responsibility in Alberta schools. Occasional Papers One. Edmonton: Alberta Teachers' Association, 1973.
- Royal Commission. *Report of the Royal Commission on the status of women in Canada*. Ottawa: Information Canada, 1970.
- Scharf, M. P. An investigation of the relationship between professional role orientation and social structure of teacher groups. Unpublished Ph.D. thesis, University of Alberta, 1967.
- Schmidt, L. C. Sex-roles and life styles of professional women. Unpublished Ph.D. thesis, University of Alberta, 1973.
- Shack, S. *The two thirds minority: Women in Canadian education*. Toronto: University of Toronto, 1973.
- Simpson, R. I., & Simpson, I. Women and bureaucracy in the semi-professions. In A. Etzioni, (Ed.), *The semi-professions and their organization*. New York: Free Press, 1969, 196-265.
- Veitch, K. E. An examination of professional role orientation among principals in Newfoundland. Unpublished M.Ed. thesis, University of Alberta, 1969.

J . W . K E H O E

University of British Columbia

Demonstrating the Relationship Between Values and Attitudes as a Means of Changing Attitudes

The purpose of the study was to test the effectiveness of two related strategies for changing a measured disposition of intolerance toward cultural diversity to one of significantly less intolerance. The strategies were generated from two sources: (1) dissonance theory of changing attitudes which suggests that if enough dissatisfaction with behaviour or beliefs exists, then change in the behaviour or beliefs will take place; and (2) Milton Rokeach's formulation that measured values are more central components of a person's make-up than measured attitudes, and values determine attitude. The first treatment attempted to determine whether a manipulated change in the relative position of three measured values related to tolerance (equality, freedom and broad-mindedness) would change the measured attitudes of students toward culturally diverse groups. The second treatment included the above sources of manipulated change and, in addition, demonstrated inconsistencies between ranks of the three relevant values and scores on the attitude scale. The treatment that was successful in changing attitudes was the second treatment. (Dr. Kehoe is Assistant Professor in the Social Studies Department, Faculty of Education, The University of British Columbia at Vancouver.)

The purpose of the study was to test the effectiveness of two related strategies for changing a disposition of intolerance toward cultural diversity to one of significantly less intolerance. The strategies were generated from two sources: Festinger's (1957, p. 13) dissonance theory and Milton Rokeach's (1968, p. 117) theoretical formulations about the relationship between values and attitudes. Dissonance theory suggests that if enough dissatisfaction with behavior or beliefs exists, then change in the behaviour or beliefs will take place. Milton Rokeach theorizes not only that measured values are more central and transcendental components of a person's make-up than measured attitudes, but also that values are determinants of attitudes. Rokeach suggests that we should be attempting to change values rather than attitudes because values provide us with a more economical tool. It is uneconomical to try to change a great many specific attitudes when all that may be needed is to enhance the importance of a small number of values related to the attitudes.

Research Questions, Sample, Design, and Treatments

To test the strategies the following research questions were investigated:

1. Did the subjects who were shown inconsistencies within their value hierarchies score significantly higher on the attitude scale on the post-test than those who did not have the inconsistency demonstrated to them?
2. Did the subjects who were shown inconsistencies within their value hierarchies, and in addition were shown inconsistencies between their values and attitudes score significantly higher on the attitude scale on the post-test than those who did not have the inconsistency demonstrated to them?

Sample

The sixty-four subjects for the study, which was completed in Surrey, British Columbia, were from two grade eleven history classes and one grade eleven law class. The subjects in the history classes were placed in the treatment groups at random. The law class was made up of students choosing it as an elective and it became the control group.

Treatments

The treatment and measurement dates are presented in Figure 1.

FIGURE 1
CHART OF TREATMENT AND MEASUREMENT DATES

	February 24	March 2	March 21
	Value Survey and Attitude Scale	Treatment	Value Survey and Attitude Scale
Control	0 ₁ , 0 ₂		0 ₃ , 0 ₄
Experimental 1	0 ₅ , 0 ₆	X ₁	0 ₇ , 0 ₈
Experimental 2	0 ₉ , 0 ₁₀	X ₁ & X ₂	0 ₁₁ , 0 ₁₂

Both experimental groups and the control group were pre-tested with the Value Survey and the attitude scale, in that order, on the same day. Approximately one week later experimental group 1 was given the following treatment.

The Value Surveys of the subjects were distributed to them. The following statements were then made to the subjects.

Look at your ranking of freedom. Now look at your ranking of equality. If your ranking of freedom is high and your ranking of equality is lower, this can be interpreted to mean that you are more interested in your own freedom than you are in freedom for other people.

Look again at your ranking of freedom. Now look at your ranking of broadminded on the second page. If your ranking of freedom is high and your ranking of broadminded is lower, this can be interpreted to mean that you want freedom for yourself but are not broadminded enough to give that same privilege to other people. If you want to have freedom, other people are going to have to be broadminded and believe in equality

so that you might enjoy the freedom you value. It is to some extent inconsistent to want such privileges for yourself but not for other people.

The subjects in experimental group 1 were then given an 11-point scale on which they were asked to indicate the degree to which they were satisfied or dissatisfied with their rankings of each of the values.

The subjects in experimental group 2 were given the same treatment as that given to experimental group 1. In addition, after the discussion of the Value Survey, their attitude booklets were returned to them with their scores on the front. The following statements were then made to experimental group 2.

Now look at the score on the front of your Information Survey. If you look at item twenty-one, I can show you how the score was calculated (item twenty-one was read to the class). If you agreed with item twenty-one, it suggests you are unwilling to support culturally diverse groups in this country. If you agreed strongly, you were given a 1; agreed moderately, a 2; agreed slightly, a 3; omitted, a 4; disagreed slightly, a 5; disagreed moderately, a 6; disagreed strongly, a 7. Look at item forty to see if you understand how the item was scored. Your score for each item was totalled to give you the score on the front of the booklet. A score under 190 is a low score. Place the numbers one to twelve down the left hand side of the front page of the Information Survey. Open your Value Survey so that you can see how *you* have ranked the values. I would now like to ask you a series of questions and I would like you to place your answers beside the numbers on the front page of your Information Survey.

1. If a person ranked broadminded high on the Value Survey, how would you expect him to score on the Information Survey, high, low, I don't know? Place your answer beside number 1.
2. If a person ranked broadminded low in the Value Survey, how would you expect him to score on the Information Survey, high, low, I don't know?
3. If a person ranked broadminded high and scored low on the Information Survey, would you conclude that his attitudes are *not* consistent with his values, yes, no, I don't know?
4. If a person ranked broadminded low and scored high on the Information survey, would you conclude that his attitudes are inconsistent with his values, yes, no, I don't know?

Now look at your own ranking of broadminded and your own score on the Information Survey.

The same four questions were asked two more times with equality and freedom substituted for broadminded in each question. After each set of four questions, the subjects were asked to compare their rank of the value with their score on the attitude scale. On completion of the questions, the students were asked to complete two separate scales indicating the degree to which they were satisfied or dissatisfied with their ranking of each of the values and the degree to which they were satisfied or dissatisfied with their score on the attitude scale.

The two experimental groups and the control group were post-tested approximately three weeks after the treatments with the Value Survey and the attitude scale. In addition the control group was asked to answer two questions after the post-testing.

- 1. Given that an individual values freedom, what is your expectation about him valuing freedom for other people as well as himself—would you expect him to want freedom for other people as well as himself, yes, no, undecided?
- 2. Given that an individual values equality, what is your expectation about him valuing equality for other people as well as himself—would you expect him to want equality for other people as well as himself, yes, no, undecided?

The control subjects were asked to place their responses to the two questions on the front of the Value Survey. The intent of the questions was to determine whether they would perceive someone as inconsistent who valued freedom and equality for himself but not for other people.

Results

An analysis of variance on the change scores of each of the four variables is reported in Table 1. No significant treatment effects were found to exist for the three values. Significant treatment effects were found to exist for the attitude scores. The Bonferroni formula was used for controls of attitude change scores (Miller, 1966, pp. 66-70). Table 2 reports the results of the contrasts.

TABLE 1
ANALYSIS OF VARIANCE ON CHANGE SCORES

Variable	Source	df	M.S.	F	p
Attitude	Treatment	2	3620.3	11.8	.0001
	Error	61	306.8		
Equality	Treatment	2	24.79	2.21	.12
	Error	61	11.21		
Freedom	Treatment	2	.59	.06	.95
	Error	61	10.52		
Broadminded	Treatment	2	29.4	1.44	.25
	Error	61	20.46		

TABLE 2
CONTRASTS FOR ATTITUDE CHANGE SCORES

Experimental 1 - Control	2.4 ± 13.68	16.08	13.68	-11.28
Experimental 2 - Control	19.7 ± 13.68	33.38	13.68	6.02*
Experimental 1 - Experimental 2	17.3 ± 13.68	30.98	13.68	3.62*

* Significant $p < .05$

An effort was made to test for comprehension of the treatment with the second experimental group. A series of twelve questions were asked throughout the treatment and the answers were written on the front of the booklet. With twenty-two subjects and twelve questions there was the possibility of 240 incorrect answers. There were twenty-six incorrect answers, most of which were to questions four, eight, and twelve. On the basis of these results, a considerable degree of comprehension was assumed to exist.

The control group of the sample was also used to investigate the question of whether subjects would perceive as inconsistent someone who valued freedom and equality for himself but not for other people. Only one subject answered one of the questions with a no. Twenty-one subjects said a person would be inconsistent if he valued freedom and equality for himself but not for other people. An interpretation that should have been a source of dissonance was one where freedom was ranked high and equality low suggesting the subject valued freedom for himself but not for others.

Conclusions and Implications

It would seem that it is possible to change the measured attitudes of students toward culturally diverse groups. The treatment given experimental group 1 was ineffective in changing values or related attitudes. The treatment given experimental group 2 was not effective in significantly changing the mean ranks of the values. It was, however, successful in changing the attitudes of the students. It would seem that the attitude scores changed because the relationship between the relevant values and the items on the attitude scale was brought to the attention of the subjects.

It has been documented by Ziegler (1967, ch. IV) that values such as those of concern to this study are not taught or discussed directly as being valuable. Nor are the students taught that the values should be accepted by a large proportion of the population. Jaros and Canon (1969), in discussing Zeigler's findings, suggest that "... the overt value role of teachers is one of abdication. Teachers, aided by the texts from which they work, strike poses of exaggerated objectivity" (pp. 236-237). A similar situation is considered to exist in Canadian schools. Democratic concepts are discussed within a chronological history of the achievement of democracy in Canada and Great Britain. There appear to be very few attempts to discuss concepts such as freedom and equality and their various meanings in the context of modern Canadian society. Nor does there appear to be much discussion of situations where values come in conflict. Saul (1960, pp. 346-347) reports that he suspects one of the possible factors which may account for frequently observed "poor" citizenship behaviour is that the individual frequently fails to discriminate particular stimulus situations as requiring action. He goes on to state that school programs should provide more explicit and varied practice in the making of such discriminations. Remmer and Radler (1957) report similar findings:

... we find that American teenagers have well-crystallized democratic attitudes with respect to religious freedom, the right to trial by jury, and equality of opportunity. But the typical teenager shows an alarming disposition to reject some democratic beliefs, to throw away some of the basic freedoms ... and to accept many authoritarian and totalitarian beliefs and values in their place. (pp. 186, 198)

Students are not given much opportunity to clarify their own value positions. For example, if a student accepts the democratic value of equality, it would appear to be useful either to help him to examine a variety of situations where he may or may not apply the value, or suggest to him that if he really accepts the value of equality as an important guiding principle in his life, then he should apply it to a wider number of situations than he does at present. Demonstrating inconsistencies between values held and attitudes or behaviour in specific and analogous situations would appear to be a useful strategy for achieving objectives like tolerant attitudes and behaviour.

References

- Festinger, L. *A theory of cognitive dissonance*. Stanford University Press, 1957.
- Jaros, D., & Canon, B. C. Transmitting basic political values: The role of the educational system. *School Review*, 1969, 5, June.
- Miller, R. G. *Simultaneous statistical inference*. New York: McGraw-Hill, 1966.
- Patterson, F. (Ed.) *The adolescent citizen*. Glencoe, Ill.: The Free Press, 1960.
- Remmer, H. H., & Radler, D. H. *The American teenager*. Indianapolis: Bobbs-Merrill Co., 1957.
- Rokeach, M. *Value survey*. Sunnyvale: Halgren Tests, 1967.
- Rokeach, M. *Beliefs, attitudes and values*. San Francisco: Jossey-Bass, 1968.
- Saul, E. V. The relevance of selected psychological data for citizenship education. In F. Patterson (Ed.), *The adolescent citizen*. Glencoe, Ill.: The Free Press, 1960.
- Zeigler, H. *The political life of American teachers*. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1967.

L. O. OLLILA

and

L. A. CHAMBERLAIN

University of Victoria

The Effect of Noise and Object on Acquisition of a Sight Vocabulary in Kindergarten Children

This study investigated the influence of classroom noise and two methods of presentation upon the rate at which kindergarten children learned four dissimilar words. A total of 120 subjects were randomly assigned to four groups of girls and four groups of boys, each containing 15 subjects. An Analysis of Variance by planned Orthogonal Comparisons revealed a significant difference ($p < .05$ level) favouring girls trained by the word-alone method under a no noise condition over girls trained by the word-object methods under a noise condition. No other significant differences were found. (Dr. Ollila is Co-ordinator of Graduate Studies, and Mr. Chamberlain is a graduate student, in the Faculty of Education, University of Victoria.)

Classroom noise and the use of prompting cues have both been of interest to educators. The purpose of this study was to investigate these two factors to discover their effects upon the rate of learning of a sight vocabulary by kindergarten children.

While researchers have long inquired into the effect of classroom noise, much concern had been generated by the recent interest in ecology. Konopa and Zimering (1972) identified noise as "our third pollution" and suggested that it has been demonstrated to be "a serious and present health hazard" (p. 116). Although educators have been exposed to claims for acoustical control in the classroom and school (McKay, 1964), researchers have reported contradictory findings.

Hartman (1946), in a review on the effects of noise upon children, stated: "Like all figure-ground phenomena, this one obeys the rules of Gestalt organization. On the whole, therefore, the more the work puts a demand on the higher mental processes, the more disturbing the noise is likely to be" (p. 159).

More recently, Kahneman (1973) suggested that the Yerkes-Dodson Law could be applied to the effect of noise on performance. Performance quality is

represented by an inverted U-shaped function of noise level (arousal). As the noise level first increases so does the quality of performance to an optimal point, and then begins to decrease with further increments of noise level.

Slater (1968) used seventh grade children to investigate three levels of noise in a classroom situation and in an experimental situation. Noise used was similar to that found in schools, but with addition of white noise (an heterogeneous mixture of sound waves extending over a wide range) in the experimental situation. Performance on the STEP Reading Test was not found to be significantly different between the groups.

Samtore (1969) reported that third grade children, especially boys, performed significantly better under noise conditions than under typical classroom conditions. Three parallel forms of The Metropolitan Reading Achievement were administered as follows: Form A—regular classroom condition, Form B—noise from a record (orchestration held at a constant level), Form C—intermittent noise from a record played at varying intensities. Variance generated by the fixed order of presentation and the differences in test forms were not accounted for in this design. Samtore concluded that a combination of factors (unspecified) and conditions produced better performance during the noise condition.

Nober (1973) administered parallel forms of the Wepman Auditory Discrimination Test to normal, speech defective, or reading retarded 5 to 8 year olds. One form was administered under a quiet situation, and the other form under a noisy situation. Both the A and B forms were randomly assigned across both noise and quiet conditions. In comparing the scores, she concluded that auditory perception was affected adversely by noise.

In these studies on the effect of noise, the measurement used was related to a task performance variable such as achievement test scores, rather than to the rate or efficiency with which learning occurred.

It has also been suggested that parents, teachers and pupils have little choice but to learn to live with noise. Children may learn to “tune out” extraneous noises by the time they enter school (Smith, 1962). However, “tuning out” has often been considered an equal problem by teachers (Wittich and Schuller, 1967).

Research in the area of prompting cues, such as the use of objects and pictures in the training of sight vocabulary, has also presented educators with contrary findings. Rohwer, Lynch, Levin and Suzaki (1967) found that an increase in learning efficiency occurred when the words were presented in sentence context, and that efficiency may be further augmented by the use of pictures. Rohwer suggested that the predicted superiority of pictures may decrease as verbal language ability increased over age. Wimer and Lamber (1959) found that undergraduates learned nonsense syllables faster when paired with objects than with words or in a mixed sequence.

Hartley (1970) reported a study on intra-list similarity and picture cues with kindergarten children learning sight words. It was found that a graphic stimulus was effective only when presented alone during study trials on minimal contrast lists. It was further suggested that if a graphic stimulus plus a context cue are to be used, they are most successful with a maximal contrast list. If a graphic stimulus alone is to be used, it is most successful with minimal contrast lists.

King and Muehl (1971) found that when pictures were presented with dissimilar words, learning was inhibited, while with similar words, learning was facilitated. Ollila and Olson (1972) using dissimilar words, found a significant difference between kindergarten boys, but not girls, favouring the word-alone method over the word-object method of presentation. In other studies with kindergarten children (Samuels, 1967; Duel, 1968), word-alone methods were found to be superior to the word-picture methods.

Samuels suggested that when pictures and words are presented together, the picture may distract or interfere with the acquisition of sight words. Similarly, Ollila and Olson (1972) reported that when prompting cues were presented, the children tended to shift or divide their attention between the words and the cue, while others attended to the cue only during the initial trials. Duel (1968) noted that those children who pointed to the cue the majority of the time learned less than those who pointed to the word or to the word plus cue.

Two summary questions appear to be pertinent: Do pictures and objects help the child make the correct association to the word, or do they cause him to not attend to the relevant features of the word, but to irrelevant or extraneous, peripheral characteristics? Does classroom noise increase, decrease or have no effect upon the rate with which sight words are learned by kindergarten children?

Samples and Procedures

120 children were selected from nine kindergartens in three Greater Victoria schools. Fifteen boys and girls were separately assigned to each of the noise conditions: noise (N), and no noise (nN), and to the two methods of presentation: word-alone (W), and word-object (W-O). Table 1 presents the seven Orthogonal Comparisons for the eight groups.

TABLE 1
PLANNED ORTHOGONAL COMPARISONS

Comparisons	Groups
1	All Boys x all Girls
2	(WnNB + WONB) x (WNB + WOnNB)
3	(WnNB) x WOnB)
4	(WNB) x (WOnNB)
5	(WnNG + WONG) x (WNG + WOnNG)
6	(WnNG) x (WONG)
7	(WNG) x (WOnNG)

Four high-frequency words (horse, ball, pencil, telephone) were printed on 5" x 8" cards using electrocet, then laminated. The objects, ball, horse, and telephone, were children's toys. An authentic pencil was used as an object for the word, pencil.

Classroom noise was recorded during an activity period with 20 children ranging in age from 6 to 12 years. A Sony cassette recorder (model TC-110) was used for both the taping and play back in the noise condition at a mean intensity of 70 dbA at four feet from the subject.

Method

Each subject met individually with the experimenter in a small medical room or remedial classroom for approximately 20 minutes on one day. All subjects received pre-testing, pre-training, and training to criterion. Criterion was defined as two successive trials within which all four stimulus words were responded to correctly. The dependent variable was the mean number of trials to criterion.

During pre-testing the subjects were told, "I would like to play a game with you to see how fast you can learn some words. If you get some wrong, it doesn't matter, but you need to listen carefully. Before we start the game, let's see if you know any of the words. What is the name of this word?" All subjects were pre-tested on all four words using a non-correction procedure. Subjects who knew one or more of the words were not included in the study. Subjects in the two W-O groups were also presented with the objects and told, "Let's see if you know any of these objects before we start the game. What is the name of this object?"

During pre-training, subjects in the two W groups were presented with each word once, and told, "This is the word _____. What is the name of this word?" Subjects in the two W-O groups were presented with word and object and told, "This is the word _____, and this is a _____. What is the name of the word?"

During training to criterion, the four stimulus words were presented in random order on each trial. WnN subjects were presented with a word, then asked, "What is the name of this word?" If the correct response was emitted, the experimenter praised the subject and proceeded to the next word. If the correct response was not emitted, the experimenter presented the subject with the word and said, "This is the word _____. What is the name of this word?" W-OnN subjects were presented with a word, then asked, "What is the name of this word?" If the correct response was emitted, the experimenter praised the subject and proceeded to the next word. If the correct response was not emitted, the experimenter presented the subject with the word and the object and said, "This is the word _____. What is the name of the word?" WN subjects received the same treatment as the WnN subjects and W-ON subjects received the same treatment as W-OnN subjects but with the tape recording of classroom noise playing. All subjects were taken to criterion or to 20 trials. Subjects who did not achieve criterion were excluded from this phase of the study, but retained for further analysis.

Results

Table 2 presents the means and standard deviations for all eight group as well as the total.

TABLE 2
MEANS AND STANDARD DEVIATIONS

Groups	\bar{X}	S.D.
WnN boys	10.00	4.80
WnN girls	7.60-	5.88
W-OnN boys	8.13	5.25
W-OnN girls	7.93	6.10
WN boys	7.80	4.80
WN girls	9.53	6.18
W-ON boys	10.20	5.71
W-ON girls	11.73-	4.71
Total	9.12	5.58

In order to test hypothesis believed important and meaningful to the interpretation of results, an Analysis of Variance by Planned Orthogonal Comparisons was applied to the data. The eight groups permitted the seven comparisons presented in Table 3 (Hays, 1963).

TABLE 3
ANALYSIS OF VARIANCE: PLANNED COMPARISONS

Source	ss	df	ms	F
Between groups	230.63	7		
Comparisons				
1	.83	1	.83	.03
2	68.27	1	68.27	2.18
3	.30	1	.30	.01
4	.83	1	.83	.03
5	13.07	1	13.07	.42
6	128.13	1	128.13	4.09*
7	19.20	1	19.20	.62
Error (within groups)	3511.73	112	31.35	
Totals	3742.36	119		

* $p < .05$

Inspection of Table 3 reveals a significant difference favouring girls trained by the word-alone method of presentation under the no noise condition over girls trained by the word-object method of presentation under the noise conditions. Remaining comparisons were not significant.

Conclusion

The rate at which kindergarten children learned four high-frequency dissimilar words was not facilitated by the word-object method of presentation. A comparison of means revealed no significant differences favouring either presentation method, and no sex differences were observed as a function of presentation method alone.

However, the combination of word-object presentation and classroom noise was found to significantly depress the performance of girls, but not that of boys. In general, the performance of boys was not affected by the presence of noise, irrespective of presentation method.

This latter finding raises an interesting question regarding sex differences in children and their ability to tolerate classroom noise. Are boys more resistant than girls to the distractions of classroom noise because of some environmental or cultural influence? Generally speaking girls are expected to behave demurely and play quietly, while boys are expected to be more aggressive in their play and hence make more noise. Perhaps in the course of their development, boys acquire a greater tolerance for noisy surroundings than girls. Or is the difference attributable to some physiological characteristic of boys, lacking in girls, that facilitates a resistance to noisy distractions. The matter warrants further investigation.

In future studies of this kind, it may be useful to employ three or more levels of noise to clarify the relationship between learning rate and noise level. As Kahneman (1973) has suggested, there may be a quadratic relationship between performance quality and noise level (arousal).

Classroom teachers should recognize that an important determinate in the learning process may be the interrelation between the noise level in the room and the particular instructional materials employed.

References

- Duel, O. K. An analysis of prompting procedures for teaching a sight vocabulary. *American Educational Research Journal*, 1968, 5(4), 675-686.
- Hartley, R. N. Effects of list types and cues on the learning word lists. *Reading Research Quarterly*, 1970, 6(1), 97-121.
- Hartman, G. W. The effects of noise on school children, an interpretive digest and recommendation. *Journal of Educational Psychology*, 1946, 37, 149-160.
- Hays, W. L. *Statistics for psychologists*. New York: Holt, Rinehart and Winston, 1963.
- Kahneman, D. *Attention and Effort*. Englewood Cliffs, N.J.: Prentice Hall, 1973.
- King, E. M., & Muehl, S. Effects of visual discrimination training on immediate and delayed word recognition in kindergarten children. *The Alberta Journal of Educational Research*, 1971, 17(2), 77-88.
- Konopa, V. A., & Zimering, S. Noise—the challenge of the future. *The Journal of School Health*, 1972, 62, 172-177.
- McKay, R. L. How to keep school noise at the right level. *The Nation's Schools*, 1964, 74, 64-67.
- Nober, L. W. Auditory discrimination and classroom noise. *The Reading Teacher*, 1973, 27(3), 288-291.
- Ollila, L. O., & Olson, J. H. The effects of learning rate of picture and realia in the presentation of words to kindergarteners. *Journal of Educational Research*, 1972, 65(7), 312-314.
- Rohwer, W. D., Lynch, S., Levin, J. R., & Suzaki, N. Pictorial and verbal factors in the efficient learning of paired associates. *Journal of Educational Psychology*, 1967, 58(5), 278-284.
- Samuels, J.S. Attentional process in reading: The effect of pictures on the acquisition of reading responses. *Journal of Educational Psychology*, 1967, 58(6), 337-342.
- Samtore, S. T. The effects of noise on a complex task. *Graduate Research in Education and Related Disciplines*, 1969, 4, 63-81.
- Slater, B. R. Effects of noise on pupil performance. *Journal of Educational Psychology*, 1968, 59, 239-243-263.
- Smith, H. R. (Ed.) Let there be noise in the classroom. *Education*, 1962, 83, 80-81.
- Wimer, C. C., & Lamber, W. E. The differential effects of word and object stimuli on the learning of paired associates. *Journal of Experimental Psychology*, 1959, 57(1), 31-36.
- Wittich, W. A., & Schuller, C. F. *Audio visual materials: Their nature and use*. (4th ed.) New York: Harper and Row, 1967.

THE JOURNAL OF
human resources

EDUCATION, MANPOWER, AND WELFARE POLICIES

Emphasizing Education—

costs of and returns from higher education and compensatory education programs;
evaluations of vocational education, training, and vocational rehabilitation;
discrimination in educational opportunities against blacks and other minority groups;
production of educational outputs;
supply of & demand for personnel in the education industry

JHR covers the broad spectrum of human resources including manpower policies, welfare policies, health care and education as they relate to the whole individual and his continuing role in society.

Regular issues of *The Journal of Human Resources* contain Articles, Review Articles, Communications, Books Received, and Book Reviews.

in health care industry—

in manpower policies—

in education—

in welfare policies—

human resources...

Special full-sized supplementary issues are published as time, material and funds become available. Previous areas covered include: "Vocational Education" and "Manpower Policies and Welfare Policies."

In 1974, The Journal of Human Resources provided the first published reports of the OEO Graduated Work Incentive Experiment, a 3-year field test of the effects on recipient families of 8 different "negative income tax" formulas.

Subscribe to The Journal of Human Resources

Founded in 1966, a quarterly publication

Editor: Stanley H. Masters

Subscription Rates: Institutions, \$20.00/year, 4 numbers; Individuals, \$10.00/year, 4 numbers (individuals must pre-pay)

*Foreign postage: 50¢ per year additional
Air Mail postage: \$12.50 per year additional*

**Send your subscriptions to:
Journals Department
University of Wisconsin Press
P. O. Box 1379
Madison, Wisconsin 53701**



The Alberta Journal of Educational Research

Volume XXI, No. 4

December, 1975

CONTENTS

Toward Teacher Education in the Year 2000	221
<i>S. C. T. Clarke and H. T. Coutts</i>	
Classroom Climate and Achievement in Secondary School Mathematics Classes	241
<i>R. O'Reilly</i>	
Classroom Personality Patterns of Teachers of Secondary Mathematics	249
<i>D. F. Robitaille</i>	
Socratic Dialogue vs. Behavioural Practice in the Development of Coping Skills	255
<i>J. P. Klein</i>	
Methodological Problems in Conservation Testing with Particular Reference to Volume Conservation	262
<i>E. D. Hobbs</i>	
An Investigation Into the Relationship of the Pausing Phenomena in Oral Reading and Reading Comprehension	278
<i>Sr. Ruth Eagan</i>	
Rural Standardisation of the Burt-Vernon Graded Word Reading Test	289
<i>G. Broadley and K. M. Broadley</i>	
Teacher Self-Actualization and Pupil Control Ideology	295
<i>L. E. Jury, D. J. Willower and W. J. DeLacy</i>	

FACULTY OF EDUCATION
The University of Alberta

S . C . T . C L A R K E

and

H . T . C O U T T S

The University of Alberta

Toward Teacher Education in the Year 2000

A modified Delphi technique and Q-sort was used to replicate and update a previous study conducted four years earlier. The views of chief administrative officers in English-language teacher education institutions in Canada about the future of teacher education were sought. The 222 statements obtained from the participants were classified and 36 representative statements about the future of teacher education were selected. The impact on education of the feature or aspects described in the statements were obtained from the Q-sort.

A feature or aspect of teacher education in the future originated by many of the participants and rated as having most impact on education was that it will be non-terminal, continuing, and lifelong. The feature judged to have next most impact on education was that teacher education in the future would feature a closer relationship between theory and practice. The study presents additional features and ratings.

There were 41 participants representing 32 of the 39 English-language teacher education institutions in Canada. (Dr. Clarke is Director of Special Sessions and Dr. Coutts is Professor Emeritus at The University of Alberta.)

Overview

The study reported here was designed in part to replicate, in part to update in 1974-75, a previous one conducted by the authors (1970-71).

In the present study a modified Delphi technique was used to elicit from the chief administrative officers of the English-language teacher education institutions in Canada their views about features or aspects of teacher education in the year 2000. The 222 statements obtained from the participants were classified, and 36 were selected to represent these features or aspects. Participants then rated these features or aspects as to probable date of acceptance, desirability, and feasibility, and were encouraged to comment on each. As a third activity, participants completed a seven-place Q-sort in which they expressed opinions concerning the impact on education of each of the 36 stated features or aspects of teacher education.

The most frequently mentioned features or aspects originated by the 45 participants were: multiple programs, lifelong or continuing involvement, humanistic concerns, increased emphasis on the school setting, a stronger or more scholarly data base, and increased selectivity. About ten percent of the comments which participants made on the 36 features or aspects of teacher education complained of ambiguity because of the use of terms such as "broader data base" and "humanistic concerns." Despite these difficulties in communication, 41 out of 45 participants completed all three parts of the study.

The five features or aspects of teacher education toward the year 2000 judged to have most impact on education, their ratings, the comments made by participants, and the number of participants who originated statements of this kind, follow:

"A recognized and accepted aspect of teacher education in the future will be that it is non-terminal, continuing, lifelong." This feature or aspect was rated as having most impact on education. Twenty-two percent of participants said it is in effect now, 22% predicted it will be before 1980, and 51% predicted it will be in effect in the majority of teacher education institutions in Canada between 1980 and 2000. Participants saw this as desirable, but moderately difficult to accomplish. Forty-two percent of the participants originated statements stating or implying this feature or aspect.

"Teacher education in the future will feature a closer relationship between theory and practice." Second in its impact on education, this feature or aspect of teacher education was judged by the participants to be in effect now or to be likely to be so before 1980. Participants saw this as very desirable, but difficult. Twelve percent of participants originated a statement along the lines of this feature or aspect.

"In the future continuing (inservice) teacher education will be equal in importance to preservice teacher education." This feature or aspect was rated as having third most impact on education. Participants were divided as to when this might be in effect in the majority of teacher education institutions in Canada. Twenty-seven percent said before 1980, 47% said between 1980 and 2000, 13% said after 2000, and 9% said never. This feature or aspect of teacher education was seen as very desirable. Participants were divided about feasibility: 22% said it was easy or very easy, 29% rated it "neutral," and 45% said it was difficult or very difficult. Forty-two percent of participants originated a statement about the importance of continuing education in the future, but not necessarily with the incisiveness of this one.

"Teacher education in the future will feature a large number of differentiated programs designed to fit the capabilities, previous experience, and career aspirations of candidates as well as their subsequent functions in newly-developed school programs." While rating this feature or aspect fourth in its impact on education, a few participants said it was in effect now, 33% said it would be before 1980, and 47% predicted it would be in a majority of the teacher education institutions in Canada between 1980 and 2000. It was rated as very desirable, but difficult. Forty-two percent of participants originated statements along the lines of this feature or aspect.

"Teacher education in the future will provide for a more rigorous screening of candidates on the basis of evaluative criteria which take into account personal as well as academic qualifications." Rating this feature or aspect fifth in its impact on education, the majority of participants saw this as in effect now or by 1980. Although 27% thought it would not be in effect in a majority of the teacher education institutions in Canada before 1980-2000, another 9% thought after 2000, and another 7% never. It was rated as very desirable and very difficult. Thirty-one percent of participants originated statements dealing with selection.

Comparisons with the previous study (1970-71) revealed an overlap of only four (about 10%) participants. The most frequently mentioned features or aspects of teacher education in the first study were also mentioned in the second. There was less correspondence in the two studies of features or aspects that had been originated by only a few participants.

Comparisons with studies of the goals of education reveal that there is no one-to-one correspondence between the goals of education as depicted in various studies and the preparatory programs of the teachers who implement those goals. The authors believe that both goals and features or aspects of the future of teacher education express a *felt need*; that is, participants do not list the most assured of likely future events nor specify as a goal that which is taken for granted as occurring.

A comparison between one study on the goals of teacher education (Coutts & Clarke, 1972) and the present study of teacher education toward the year 2000 showed a high degree of congruence.

Population

The chief administrator in each English-language teacher education institution in Canada was invited to participate, and in the event that the institution enrolled more than 1,000 students, to name an additional person for each 1,000 students or fraction thereof.

Of the chief administrators of the 39 institutions contacted, 32 agreed to participate. Of the 20 nominees of the chief administrators who agreed to participate, 4 were assistants, associates, deputies, or equivalent; 14 were department chairmen or heads, directors, coordinators, or equivalent; and 2 had no administrative title. Seven institutions originally contacted were not represented in the study.

The views of teacher education toward the year 2000 as expressed in this study are the views of the chief administrators and their immediate colleagues in 32 out of 39 of the English-language teacher education institutions in Canada. The 32 institutions included 91% of the students enrolled.

Statements about the future of teacher education

Each of the 52 persons who agreed to participate was sent an offprint of the *Journal of Teacher Education* article which described the 1970-71 Clarke-Coutts study and a single sheet with these instructions: "Thinking about the future as we approach the 21st century, what will teacher education be like? You may wish to precede each statement you make with the stem 'I believe that teacher education in the future will . . .'" The sheet bore the title of this study and a subtitle "STATEMENTS ABOUT

TEACHER EDUCATION IN THE FUTURE." Most of the participants (89%) produced 4, 5, or 6 statements. The range was 3 to 10.

The 222 statements obtained were classified, as shown in Table 1, in order to facilitate the educing of 36 representative items. Since the basic purpose of the classification was to facilitate the educing and phrasing of representative statements, no special efforts were made to refine or validate the classes. In some instances, there was little difficulty about classification, e.g., that in the future there would be more emphasis on selection: initially, throughout preservice teacher education, at the end before placement, etc. In other instances, the classification was difficult, e.g., as between humanistic concerns and stress on personal development. Some statements made by participants contained more than one idea. It was decided that the core or central idea, as judged by the authors, should be used as a single statement because this would seem to be a more reasonable representation of the views of the different participants than would result from breaking up omnibus statements into two, five, or possibly ten ideas included. This decision made classification more difficult.

For the reasons stated above undue importance should not be placed on the classes and the number of statements presented in Table 1. It is the opinion of the authors, however, that the classification is sufficiently valid to warrant the presentation of the classes and the frequency of statements therein as a first approximation to the future of teacher education as seen by the participants. The reader is reminded again that these participants were the chief administrators in most of the English-language teacher education institutions in Canada.

There were 45 participants who submitted statements. The commonality of statements was not great: at most 19 out of 45 participants (or 42%) originated a statement of the same classification. Six classes included 105 (53%) of the 199 statements classified as other than miscellaneous. Viewed in this way, one can detect a preoccupation with multiple programs, lifelong teacher education, humanistic concerns, school setting, a stronger data base, and selection, in the statements which came to the minds of the chief administrators as they contemplated teacher education about the year 2000.

Representative statements

The task of educing a set of statements about teacher education in the future representative of those originated by the participants drew heavily on the scholarship, impartiality, and command of English usage of the authors. The classification, made for this purpose, simplified the task. Admittedly statements should be clear and unambiguous, should each present one and only one idea, and should be representative of the original set. Because participants would be asked to rate and comment on the statements in Part II and to sort them in Part III of the study, it was decided arbitrarily to limit the number to 36. The statements in each classification were examined and the number of items necessary to represent the ideas within various categories was determined; for example, the first three statements are based on the "Multiple Programs" classification. It should be emphasized that the procedure used focussed on majority views;

TABLE 1
CLASSIFICATION OF STATEMENTS ABOUT THE FUTURE OF
TEACHER EDUCATION

Classification	Number of Statements
1. MULTIPLE PROGRAMS: individualized (8), differentiated but not individualized (11)	19
2. LIFELONG OR CONTINUING: using the term "inservice" (10), using the term "continuing" (9)	19
3. HUMANISTIC: human relations, interpersonal skills, communication skills, commitment to social needs	19
4. INCREASED EMPHASIS ON SCHOOL SETTING: internship (10), field experience (6), apprenticeship (2)	18
5. STRONGER OR MORE SCHOLARLY DATA BASE: more knowledge about human behaviour	16
6. MORE SELECTIVE	14
7. ONE OUTCOME STRESSED: diagnostician, remediator, curriculum developer, tutor, team member, facilitator of learning, etc.	11
8. MORE SPECIALIZATION	9
9. CONTROL AND RESPONSIBILITY INCREASINGLY SHARED: with teachers' organizations, school boards, government, etc.	9
10. CONTROL AND RESPONSIBILITY TO MOVE OUT OF THE HANDS OF THE UNIVERSITY	7
11. COMPETENCY BASED EMPHASIS	7
12. TEACHERS TO BE PREPARED FOR ROLE ADDITIONAL TO TEACHING GRADES 1-12 (13)	7
13. INCREASED EMPHASIS ON THE COMMUNITY SETTING	7
14. INCREASED TOTAL-UNIVERSITY INVOLVEMENT	6
15. GREATER RESEARCH ORIENTATION	6
16. INTERDISCIPLINARY EMPHASIS, INTEGRATED, LESS FRAGMENTED	5
17. INCREASED EMPHASIS ON PERSONAL DEVELOPMENT	5
18. UNIFIED PROGRAMS: marriage of theory and practice	5
19. MORE EMPHASIS ON PREPARING GENERALISTS	4
20. GREATER EMPHASIS ON VALUES	3
21. LONGER	3
22. MISCELLANEOUS: not classifiable in above categories, largely because they were statements about education	23
TOTAL	222

highly individual views could be lost within a classification or because of the 36 statement limitation.

The educed statements were listed in Part II of the study in order as derived from Table 1 with those representing frequent mention coming first, and those representing infrequent mention coming last. Two items which kept reappearing as asides in statements which were classified elsewhere were included (numbers 33 and 34): teacher education in the future will be about the same as it is now, and teacher education in the future will bear little resemblance to what it is now.

Comments on statements

Part II invited participants, as they rated the items, "to comment on your decision in the space provided." Only four of the participants made no comments. At the other extreme, six participants commented on every item. The average number of comments was 16.5.

There were 73 comments classified as Communication, or an average of 1.7 per participant; 129 classified as Explanation of Rating, or an average of 2.9 per participant; and 541 classified as Comments on Substance for an average of 12.0 per participant. The difficulty in communication is illustrated in item 10, "Teacher education in the future will feature a much greater emphasis on humanistic concerns," where 10 participants (22%) found the term *humanistic concerns* too vague. An examination of the 36 items reveals some where the expectation of possible ambiguity is high, but where none was mentioned; for example, item 22, "Teacher education in the future will involve the community and community resources to a much greater extent." Insofar as participants experienced difficulty in communication, their willingness to rate the items was reduced.

The comments which explained the participant's ratings are self-explanatory. Those classified as substantive were observations about the cost, resources required, social pressures, agreed-upon goals, and trends which the item generated in the minds of the participants.

The comments reveal that lack of clarity of meaning casts doubt on the finding associated with items 10, 12, and 31. Yet such is the idiosyncratic nature of communication that many participants appeared to experience no difficulty with these items. One of the most interesting findings is the small variation in the number of comments per item. A careful perusal of the comments does reveal an individual interpretation on many of the items which may mean that there is less true agreement than the findings show.

Ratings

In Part II of the study each participant was asked to rate the probable date of acceptance, desirability, and feasibility of each item. Table 2 present the items, the results in numbers and a descriptive analysis. For the first item, "The basic program of teacher education in the future will vary in length for different candidates," 4 participants (9%) indicated that this is in effect now, 15 (33%) that it will be before 1980, 13 (29%) that it will happen between 1980 and 2000, 8 (18%) that it will occur after 2000, and 3 (5%) that this statement about the future of teacher education would never hold. There were two (4%) participants who did not rate the probable date of acceptance, for a total of 45 participants.

The analysis of the comments in the preceding section revealed that for many of the participants items 10, 12, and 31 were not clear. An inspection of the "not rated" column of Table 2 shows that for item 10, 8 participants (18%) did not rate desirability or feasibility, and 9 participants (20%) did not rate probable date of acceptance. For item 12, the numbers who did not rate the corresponding parts were 15, 15, and 13 respectively, and for item 31 were 8, 7, and 7 respectively. Obviously, one reaction of participants was to decline to rate a statement which they did not understand. Two other items (33 and 34) frequently were not rated because they did not fit the

rating system, in that one stated that teacher education in the future would be about the same as it is now, while the other stated that it would bear little resemblance to what exists in 1975. Three other items were not rated by a number of participants: 18 and 19 which dealt with who, in the future, would control teacher education, and 26 which predicted emphasis on preparation of generalists.

Using a simple weighting where very desirable is assigned a weighting of 5, desirable 4, neutral 3, undesirable 2, and very undesirable 1, the features or aspects of teacher education found most desirable were as follows in rank order of the combined weightings for "very desirable" and "desirable": #7 (lifelong or continuing involvement), #27 (closer relationship between theory and practice), #9 (emphasis on human relations), #3 (differentiated programs), #29 (interdisciplinary programs and procedures) and #6 (greater proportion to resources to inservice) tied, #15 (stress on personal development), #16 (total university commitment), #22 (involvement of the community), and #2 (individualization).

Using a similar weighting system with respect to difficulty, rank order in decreasing difficulty is as follows: #23 (teaching practice linked to apprenticeship of health service neophytes), #14 (selection), #16 (total university commitment), #3 (differentiated programs), #15 (stress on personal development), #27 (closer relationship between theory and practice), #28 (knowledge focussed on applied skills), #13 (emphasis on research) and #29 (interdisciplinary programs and practices) tied, and #2 (individualization).

There is a heavy overlap between the ten most desirable and the ten most difficult items, #27 (closer relationship between theory and practice), #3 (differentiated programs), #29 (interdisciplinary programs and practices), #2 (individualization), #15 (stress on personal development), and #16 (total university commitment) being on both lists.

At the other end of the scale, the five features or aspects judged as least desirable were: #33 (teacher education will remain as it is now), #19 (move out of the hands of the universities), #17 (joint control by teacher organizations and teacher education) and #18 (joint control: universities, professional organizations, employing agencies, government) tied, and #20 (centre on the development of competencies).

The five rated as most easily accomplished were: #9 (emphasis on human relationships), #16 (a greater proportion of resources to inservice), #1 (varying length of program), #4 (more time in schools), and #22 (community involvement).

We can now identify two items deemed to be desirable (in the top ten) and easy (in the top 5): #9 (emphasis on human relations) and #6 (a greater proportion of resources to inservice).

The rating of probable date of acceptance permitted participants to indicate disagreement with a feature or aspect which had been originated by other participants. In descending order, the following features or aspects were rated as "never" for probable date of acceptance: #19 (teacher education will move out of the hands of the universities), #33 (teacher education will remain as it is now), #18 (joint control: universities, professional organizations, employing agencies, government), #17 (joint control by teacher organizations and teacher education institutions), #30 (longer), #25

Table 2
Ratings of Probable Date of Acceptance, Desirability, Feasibility, and of Impact on Education
of Thirty-six Features or Aspects of Teacher Education toward the Year 2,000

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated		Impact on Education							Wtd Total	Rank Order					
N = Number Wtd = Weighted															Q-Sort Frequency MOST...to....LEAST													
															Weighting													
1. The basic program of teacher education in the future will vary in length for different candidates.	N	4	15	13	8	3	18	17	7	2	0	3	16	6	14	5	2	1	1	1	1	3	16	13	4	3	142	31
	%	9	33	29	18	7	40	38	16	4	0	7	36	13	31	11	4	2	2									
2. Teacher education in the future will be much more individualized, with candi- dates having much more say in the design of their own programs.	N	2	13	28	1	0	18	23	2	1	0	1	12	5	26	0	1	1	1	1	2	12	21	2	3	0	175	11
	%	4	29	62	2	0	40	51	4	2	0	2	27	11	58	0	2	2	2									
3. Teacher education in the future will feature a large number of different- iated programs designed to fit the capabilities, previous experience, and career aspirations of candidates as well as their subsequent functions in newly- developed school programs.	N	2	15	21	5	2	21	21	1	0	1	1	7	4	27	4	0	1	2	5	6	5	20	5	0	0	191	4
	%	4	33	47	11	4	47	47	2	0	2	2	16	9	60	9	0	2	4									
4. In the future candidates will spend a notably greater portion of the time devoted to their programs in schools.	N	9	25	11	0	0	20	17	5	3	0	2	17	14	11	1	0	0	0	1	4	12	15	7	2	0	176	9
	%	20	56	24	0	0	44	38	11	7	0	4	38	31	24	2	0	0	0									
5. Teacher education in the future will include a supervised internship with a total duration of approximately a year.	N	5	12	21	3	1	20	15	7	2	0	2	6	15	19	1	3	1	2	0	4	12	14	9	2	0	169	19
	%	11	27	47	7	2	44	33	16	4	0	4	13	33	42	2	7	2	4									
6. Teacher education in the future will devote a greater proportion of its total resources to inservice education.	N	7	19	15	1	1	27	13	3	1	0	1	19	10	12	2	2	1	1	1	5	16	15	2	2	0	187	6
	%	16	42	33	2	2	60	29	7	2	0	2	42	22	27	4	4	2	2									

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order											
	Now		Before 1980		1980-2000		After 2000		Never		Very Desirable		Desirable		Neutral		Undesirable		Very Undesirable		Very Easy		Easy		Neutral		Difficult		Very Difficult		Date		Desirability		Feasibility
N = Number Wtd = Weighted																																			
7. A recognized and accepted aspect of teacher education in the future will be that it is non-terminal, continuing, lifelong.	N 10 % 23	10 23 22 51	23 51 4 0	2 4 0 0	0 0 0 0	37 82 7 16 0 0	7 16 0 0	0 0 2 2	1 2 0 0	0 0 0 0	2 4 22 33	10 27 12 27 15 33	12 27 27 33	15 33	5 11	5 11	0 0 0 0 2 2	1 0 0 2	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	3 9 11 14 1 0	197	1	
8. In the future continuing (inservice) teacher education will be equal in importance to preservice teacher education.	N 1 % 2	12 21 27 47	21 47 13 9	6 4 13 9	2 4 0 0	24 53 11 24 13 24	11 24 24 53	6 13 13 29	2 2 4 4	0 0 0 0	1 2 20 29	9 20 13 29 13 29	13 29 29 33	13 29	7 16 16 24	7 16 16 24	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	2 2 2 4 4 4	193	3	
9. In the future there will be an increased emphasis on human relations in preparing candidates for teaching.	N 10 % 22	19 42 15 33	15 33 0 2	0 2 2 2	1 1 2 2	28 62 14 31 3 7	14 31 3 7	3 7 7 0	0 0 0 0	0 0 0 0	3 7 40 29	18 40 6 13 13 29	6 13 13 29	13 29	5 11	5 11	0 0 0 0 0 0	0 0 0 0 0 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	2 8 7 16 2 0	183	8		
10. Teacher education in the future will feature a much greater emphasis on humanistic concerns.	N 4 % 9	14 31 15 33	15 33 0 7	0 7 3 7	3 7 7 7	21 47 13 29 7 29	13 29 29 33	3 7 7 0	0 0 0 0	0 0 0 0	2 4 4 4	11 24 24 53	11 24 24 53	18 33	5 11	5 11	8 8 18 18 18 18	8 8 18 18 18 18	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	3 1 7 23 3 1	173	15		
11. Teacher education in the future will be based on a hard central core of knowledge (nature of development, learning theory, communication and social interaction theory, educational design, etc.) and professional skills, with a corresponding diminution of "conventional wisdom."	N 3 % 7	10 22 15 33	15 33 22 11	10 22 11 11	5 5 11 11	17 38 14 31 7 16	14 31 31 16	7 16 16 11	5 0 11 0	0 0 0 0	1 2 20 27	9 20 27 33	12 27 24 33	11 24	8 18	8 18	2 4 4 9 4 9	2 4 4 9 4 9	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	1 7 8 10 1 7	174	13		
12. Teacher education in the future will rest on a broader data base and will demand that candidates possess a mastery of this base.	N 1 % 2	9 20 13 29	13 29 4 9	4 9 11 11	5 5 11 11	7 16 16 38 11 24	16 38 38 11	5 11 11 24	1 2 2 0	0 0 0 0	0 0 11 29	5 13 11 29	13 29	11 24	1 2	1 2	15 15 33 33 33 33	15 15 33 33 33 33	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	1 3 4 14 1 3	154	25	

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order			
N = Number Wtd = Weighted																Q-Sort Frequency MOST....to....LEAST											
																Weighting											
13. Teacher education in the future will feature increased emphasis on research in teaching and teacher education.	N 3	13	18	8	1	19	17	7	0	0	1	5	10	23	4	2	2	2	2	4	9	17	7	2	0	176	10
14. Teacher education in the future will provide for a more rigorous screen- ing of candidates on the basis of evaluative criteria which take into account personal as well as academic qualifications.	N 5	18	12	4	4	30	6	3	2	2	1	4	2	15	21	2	2	2	4	5	11	15	6	0	0	191	5
15. Teacher education in the future will feature much greater concern for, and devote much more time to, the personal development of candidates.	N 2	15	20	5	1	22	19	1	1	0	1	5	7	23	7	2	2	2	1	1	12	19	4	3	1	168	21
16. Teacher education in the future will involve much more concern and com- mitment by the total university.	N 2	6	20	10	5	26	13	4	1	0	1	2	5	19	16	2	1	2	0	2	6	15	10	7	1	147	28
17. Teacher education in the future will be under the joint control of teachers' organizations and teacher education institutions.	N 2	2	20	4	14	5	11	14	7	8	1	5	15	11	10	3	0	3	1	2	0	11	12	11	4	125	33
18. Teacher education in the future will be controlled jointly by universities, professional organizations, employing agencies, and government, in that order	N 7	3	6	6	17	6	10	9	7	8	3	6	8	8	13	6	5	7	0	0	3	13	16	9	0	123	35
19. In the future, control of teacher education will move out of the hands of the universities.	N 2	5	3	6	21	0	4	12	10	15	2	6	14	5	10	8	4	8	1	0	3	11	9	14	3	124	34

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																															
N = Number Wtd = Weighted																Q-Sort Frequency MOST...to....LEAST																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																							
	Now	Before 1980	1980-2000	After 2000	Never	Very Desirable	Desirable	Neutral	Undesirable	Very Undesirable	Very Easy	Easy	Neutral	Difficult	Very Difficult	Date	Desirability	Feasibility	Weighting																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
																			7	6	5	4			3	2	1																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated		Impact on Education							Wtd Total	Rank Order													
N = Number Wtd = Weighted	Now				Never				Very Desirable				Desirable				Neutral				Undesirable				Very Undesirable				Q-Sort Frequency MOST....to....LEAST							
	Before 1980				1980-2000				After 2000																				Weighting							
26. Teacher education in the future will emphasize the preparation of general- ists more than specialists.	N	8	8	9	2	12	7	18	8	3	4	4	14	13	4	3	6	5	7	0	3	8	11	11	7	1	150	27								
27. Teacher education in the future will feature a closer relationship between theory and practice.	N	5	21	15	2	1	31	12	1	0	0	3	8	5	17	10	1	1	2	4	5	10	21	1	0	0	195	2								
28. Teacher education in the future will feature an integration of knowledge from whatever source to focus on applied skills and competencies req- uired by the effective teacher, rather than a curriculum based on tradit- ional disciplines and courses.	N	3	13	13	8	3	18	15	3	3	2	0	5	9	19	8	5	4	4	1	4	7	21	7	1	0	173	16								
29. Teacher education in the future will be increasingly interdisciplinary with contributions from a wide range of specialists in appropriate dis- ciplines being focussed on problems specific to teaching.	N	0	13	23	7	0	19	23	1	0	0	0	4	12	18	8	2	2	3	1	3	13	17	7	0	0	169	20								
30. Teacher education in the future, while enabling candidates to teach after the completion of a basic pro- gram of teacher education at degree level, will increasingly involve 5 to 8 years of initial preparation.	N	1	4	13	11	13	4	16	9	10	3	3	7	9	15	5	3	3	6	0	2	3	15	16	4	1	144	30								
31. Teacher education in the future will increasingly prepare candidates to be facilitators of learning.	N	9	12	15	2	0	19	17	1	0	0	1	11	4	14	8	7	8	7	2	8	8	14	8	1	0	184	7								

Features or Aspects of Teacher Education toward the Year 2,000	Probable Date of Occurrence				Desirability				Feasibility				Not Rated			Impact on Education							Wtd Total	Rank Order
N = Number Wtd = Weighted																Q-Sort Frequency MOST...to....LEAST								
																Weighting								
	Now	Before 1980	1980-2000	After 2000	Never	Very Desirable	Desirable	Neutral	Undesirable	Very Undesirable	Very Easy	Easy	Neutral	Difficult	Very Difficult	Date	Desirability	Feasibility	7	6	5	4		
32. Teacher education in the future will feature the preparation and facilitate the placement of candidates consistent with their personal characteristics.	N 0	6 13	17 38	9 20	8 18	13 29	15 33	6 13	3 7	3 0	3 7	12 27	10 22	15 33	5 11	5 11	0 1	12 19	5 4	0 0	165	23		
33. Teacher education in the future will be just about the same as it was in 1975.	N 3	5 11	1 2	0 0	19 42	0 0	0 11	5 38	17 38	6 13	7 16	8 18	4 9	3 7	17 38	6 13	1 1	0 6	4 7	22	85	36		
34. Teacher education in the future will bear little resemblance to teacher education in 1975.	N 1	1 2	8 18	17 38	3 7	7 16	17 38	7 16	4 9	1 2	0 4	2 22	10 20	9 24	11 24	9 20	1 1	5 13	11 8	2 2	141	32		
35. Teacher education in the future will involve increased use of technology such as computer assisted instruction, instructional packages, videotaped materials, etc.	N 5	18 40	16 36	6 13	0 0	7 16	27 60	9 20	2 4	0 2	1 29	13 24	11 40	18 40	2 4	0 0	0 0	2 8	14 14	3 0	156	24		
36. Teacher education in the future will feature an increasing emphasis on controlled clinical experiences (micro-teaching, simulation, tutoring, highly structured and analyzed mini teaching episodes, and the like) in contrast with traditional and/or extended practicums.	N 5	9 11	17 20	2 38	7 10	13 29	16 36	6 13	2 4	5 4	2 11	5 36	17 38	1 2	5 7	3 9	4 9	17 6	3 3	0 0	171	18		

(more specialized) and #26 (prepare generalists) tied, #32 (preparation and placement consistent with personality), #20 (centre on the development of competencies) and #36 (increased emphasis on clinical experiences). The number of persons who rated the feature or aspect of teacher education as never likely to occur decreased from 21 who so rated item 19 to 7 who so rated item 36.

As might be expected, features or aspects of teacher education rated as least desirable are also rated as never likely to be accepted or occur. All five of the least desirable items (33, 19, 17, 18, and 20) are found in the ten rated least likely to be accepted or occur.

Further study of Table 2 reveals that there was, indeed, a definite future cast in the 36 statements. Very few participants indicated that the feature or aspect of teacher education described was in effect now. For two, #7 (lifelong or continuing) and #9 (emphasis on human relations) 10 participants (22%) indicated that this feature was in effect now. All other items had a lower ratings.

The analysis of participant's ratings in Table 2 summarized the major reactions to the features or aspects of teacher education delineated by the 36 items according to the preceding analysis.

Impact on education

In Part III of the study, a specially prepared Q-sort requested participants to select the one feature or aspect of teacher education with most impact on education, the three with next most impact, then the seven, fourteen, seven, three, and finally the one with least impact on education. Forty-one of the 45 participants who had completed Parts I and II completed Part III.

Participants observed that it was difficult to include two statements "that teacher education in the future would be just about the same as now" or "will bear little resemblance to teacher education in 1975." The items previously noted as not being clear might have caused difficulty in the Q-sort, but a perusal of Table 2 shows that many participants were able to sort items number 10, 12, and 31. These items, however, were not ranked high in their impact on education.

An examination of Table 2 also reveals that item 7, "a recognized and accepted aspect of teacher education in the future will be that it is non-terminal, continuing, lifelong," was rated the feature or aspect of teacher which would have most impact on education toward the year 2000. Three participants rated it as the one feature or aspect with most impact, 9 participants rated it as the feature or aspect with next most impact, and 11 participants rated it as the feature or aspect with next most impact. The weighting scheme listed provided a weighted total of 3x7, 9x6, 11x5, and so on for a total of 197 (see Table 2).

The ten features or aspects rated as having most impact on teacher education were: lifelong or continuing involvement; closer relationship between theory and practice; inservice teacher education equal in importance to preservice; differentiated programs; more rigorous screening; a greater proportion of total resources to inservice education; candidates prepared to be facilitators of learning; increased emphasis on human relations; candidates spending more time in schools, and increased

emphasis on research in teaching and teacher education. The ten features or aspects with least impact on teacher education were: teacher education in the future will be just about the same as it was in 1975; control will be jointly by universities, professional organizations, employing agencies, and government in that order; control will move out of the hands of the universities; control will be jointly between teacher organizations and teacher education institutions; teacher education will bear little resemblance to teacher education in 1975; the basic programs will vary in length for different candidates; will involve 5 to 8 years of initial preparation; will be much more specialized; will involve more concern and commitment by the total university; and will emphasis the preparation of generalists.

The diversity of opinion among participants is apparent. Only 13 of the 36 features or aspects were not rated by any participant as having “most impact” on education toward the year 2000, and 24 of the 36 were not rated as having “least impact.” Inspection reveals five items (1, 15, 16, 17, and 19) which were rated by participants in every possible way, from most impact to least impact on education. Despite this diversity, the agreement among participants with respect to the features or aspects with most impact is impressive, and perhaps even more so with respect to least impact.

Comparisons

Since this study was designed to replicate, at least in part, a 1970-71 study on the same topic, the first comparison will be with this earlier study.

The basis for choosing participants was identical in each study: the chief administrative officer, and in the event that the institution enrolled more than one thousand students, an additional person for each 1000 students or fraction thereof, in the English-language teacher education institutions in Canada. Table 3 shows the comparisons.

TABLE 3
COMPARISONS OF PARTICIPANTS

	Institutions	Institutions Represented	Individuals Contacted	Participants Who Completed all Parts
1970-71	41	27	55	40
1974-75	39	32	61	41

The overlap of participants after four years was four persons of whom three were chief administrators and one acting for the administrator who was on leave. Turnover is high, holding power is low, drop-outs abound among chief administrators and their designees! From the point of view of the two studies, with a ten percent overlap in participants, there are basically two different sets of participants.

The overlap in statements about the future of teacher education was considerable. In the first study there were 38 statements of features or aspects. While wording varied there were 21 statements common to the two studies. These common statements express the view that there will be a longer

period of preservice preparation; extensive use of internship; many teachers prepared for other than K-12; teacher education just about the same as it is now; teacher education that is lifelong or continuing; more stress on selection with reference to speech, English usage, mental health, and human relations; shared control with teachers' organizations concerning goals, curriculum, recruitment, selection and certification; more specialization; individualization; differentiated programs; more attention to competency based approaches; stress on ethics and values; preparation in the use of educational technology; individualization to the point of no commonality; preparation of teachers as facilitators of learning. It should be noted that in two instances (selection and shared control) the earlier study had several statements expressing different features or aspects while the present study has only one.

Features or aspects contained in the earlier study but not found in the present one follow: Teacher education will be the responsibility of universities; there will be no *institutions* devoted to preparing teachers; teacher education will be involved with constant or periodic re-evaluation of teachers who will have to requalify to retain certification; certification specific to the area of specialization; two major branches of teacher education: educational technology and educational psychology; preparation in working as a leader or member of a group; half "common core" and half specific to specialization re function, level, staff differentiation; general education and subject specialization relegated to the background by psychological, sociological, cultural values, guidance, and similar studies; general methods related to the cognitive disciplines replacing specialized methods; lectures to disappear; professors of education to spend as much time in the field as they do in the university; the professor of education to remain only as a research scholar; disillusionment with innovation and change in teacher education; candidates specialized both in individualization and in group processes; teacher education to be involved in making education an effective force for reducing social and other inequalities.

The features or aspects found in both studies are the ones mentioned most frequently in the initiation of items by participants (Part I), while the differences between the two lists occur in the features or aspects originated by fewer participants.

Some of the features or aspects found in the second study which were not specifically mentioned in the first are: emphasis on humanistic concerns; stronger data base (more scientific); emphasis on research in teacher education; personal development; concern and commitment by the total university; control of teacher education to move out of the hands of the universities; community involvement; a closer relationship between theory and practice; interdisciplinary involvement; preparation and placement consistent with personal characteristics; increased emphasis on controlled clinical experience.

Considering that there was a time lapse and considering that the second study substantially (90%) involved different people, the overlap between the features or aspects mentioned is substantial. Views of the future may change, but in the short run, they retain a considerable stability.

It is not possible to generalize on the basis of the participants common to the two studies, but some interesting comparisons are possible. First,

they tend to be no different than other participants in the study, their initial statements and their reactions to the items in Part II reflecting this. Second, they have generally moved with changes that have occurred since 1970-71 and are aware of current and anticipated thinking and practice with respect to teacher education. Third, they seem more concerned with broad aspects of teacher education than with narrower ones. Fourth, they are consistently in favour of quality programs, which, besides assuming a strong base in knowledge and educational theory and practice, are designed to meet changing times and conditions. Fifth, they recognize the importance of the personal development of the neophyte teacher. In general, they would rate high those items so rated by the participants in this study.

The first study did not ask participants to rate feasibility, desirability, or impact on education, so further comparisons cannot be made on these matters. An interesting task in the first study was that participants were asked to rate when each of the features or aspects of teacher education would be descriptive of teacher education in their own institutions. The finding was that the features or aspects would occur at later dates in their own institutions than in teacher education in general.

A second comparison which can be made is to examine the consistency between forecasts of the future of teacher education, and what people consider to be the goals of education. Lauwerys (1973) stated the conclusions of a 1972 cross-Canada survey of the purposes of education. Included were: diversity is a good thing, the prime function of schools is to convey knowledge and to develop reasoning power, education should aim at helping individuals to enjoy life more richly, e.g., through the arts and by fostering creativity, schools should help children toward getting a job, but, what is even more important, should help them attain self-satisfaction. One can ask whether the future of teacher education should reflect these goals, and whether in our study it did.

It is the view of the authors that forecasts of the future and statements of goals both express participants' *felt needs*. For example, in 1970 not all teacher education in Canada was located in universities or university related institutions. Participants in the study originated this feature or aspect as applying to the future of teacher education toward the year 2000. In 1975, when in fact teacher education in Canada was located in universities, the *felt need* had disappeared and not one of the participants proposed this as a feature of teacher education in the future. If the *felt needs* view is correct, then a one-to-one correspondence between the future of teacher education and the goals of education can be expected only to the extent that the felt needs apply to both.

Perhaps the clearest example is that the persons who responded to the C.E.A. study of Canadian education (Lauwerys, 1973) considered that a prime function of the schools was to convey knowledge and develop reasoning power. No feature or aspect of teacher education in the present study relates directly to this prime goal, presumably because the participants felt that teachers were already, and would continue to be, prepared to promote this goal.

There is consistency, however, between the goals of education and the future of teacher education. For example, each emphasizes diversity, concern for individual development, good human relations, and general

humanistic concerns. These are felt needs; that is, areas for increased emphasis or concern. But there is no one-to-one correspondence between the goals of education (C.E.A.) and the future of teacher education (present study).

Another study, "1,000,000 Opinions: A Report of the Social Values Survey" which Chalmers (1972) conducted in Greater Montreal for the Protestant School Board, found that parents ranked educational goals in this order: vocational preparation, developing self-discipline, helping students understand themselves and others, moral and ethical values, and encouraging originality and creativity. The present study does not mention teacher education to enable teachers to prepare students for the world of work or to get a job, presumably because of the *felt needs* explanation. The other goals, either directly or by implication, are consistent with the preparation of teachers shown in this study.

At the opposite side of Canada, in a study by Coutts, Clarke and Horowitz (1974), in which the goals of education were originated by the panelists, the ten most frequently mentioned were personal development, critical thinking, vocational preparation, universality of educational opportunity, imparting knowledge, constructive use of leisure time, moral and ethical values, lifelong or continuing involvement, ability to cope with change, and developing attitudes favourable to social survival. The study referred to, centered in the province of Alberta, sought the views of a number of groups. Again we find many of the goals of education not specifically covered by the features or aspects of teacher education toward the year 2000 as identified in the present study.

A study by Coutts and Clarke (1972) of the goals of teacher education, also centered in Alberta, which permitted participants to originate the goals, showed that the ten most frequently mentioned were change, values, methods, human relations, inservice education, individualized instruction, selection, facilitation of learning, community involvement, and the learning process. There is much more congruence between these findings and the features or aspects of teacher education toward the year 2000. The notable exception is the strong emphasis on change in the study cited, an emphasis which may have been produced by the procedures used.

It would seem, then, that there is no one-to-one relationship between what various groups see as the goals of education, and our panelists see as the future of teacher education. The authors attribute this in large part to the *felt needs* explanation. This, however, is a matter for some concern and for further study. In the one study cited there was good correspondence between the goals of teacher education as seen by various groups in one province, and the future of teacher education delineated by the participants in the present study.

Discussion

The roles people play and the social setting at a particular time are likely to affect views about the future. This study found that when role was held reasonably constant (the chief administrator in English-language teacher education institutions in Canada) despite the fact that the personnel involved were almost completely different, the view of teacher

education toward the year 2000 remained remarkably constant over a four to five year period. It also identified examples of the social setting at a particular time affecting views about the future, as illustrated in the dropping of the forecast that teacher education would be university based.

The authors have put forward a *felt needs* theory about the generation of statements of goals, or of forecasts about the future in human affairs such as education. The commonplace, the inevitable, the certain, that which is being done and not attracting attention, are the matters which are ignored. The uncommon, the less likely, the uncertain, that which should be done or should be done better, the gap between what *is* and what *ought to be* generate the statements about goals and about the future.

Summary and Conclusions

1. The chief administrative officers in English-language teacher education institutions in Canada when asked to make statements about teacher education toward the year 2000 mentioned most frequently multiple programs, lifelong or continuing involvement, humanistic concerns, an increased emphasis on the school setting, and a stronger data base (more scholarly).

2. The 222 statements originated by participants were classified and 36 representative statements were selected to represent the views expressed. Communication difficulties were illustrated by three of the statements regarded as vague by some participants, e.g., humanistic concerns, facilitators of learning, broader data base.

3. Participants rated the 36 representative statements as to the time when the feature or aspect will be in effect in the majority of teacher education institutions in Canada, its desirability, its feasibility. Many commented on the substance of the statement, explained the reason for their rating, or complained about the communication factor.

4. The features or aspects of teacher education considered to have most impact on education were that in the future it will be lifelong or continuing, will exhibit a closer relationship between theory and practice, will stress in-service teacher education as equal in importance to preservice, will provide differentiated programs, and will involve more rigorous screening.

5. In general, the features or aspects of teacher education toward the year 2000 deemed to have most impact on education were seen as desirable or very desirable, difficult or very difficult. The road is hard!

6. There is substantial agreement with the view of teacher education toward the year 2000 as seen by participants in an earlier (1970-71) study. The participants were substantially different persons holding the same positions.

7. There is no one-to-one correspondence between teacher education toward the year 2000 as seen by the participants in this study and the goals of education as outlined in various studies. This is explained, in part at least, by differences in the *felt needs* of participants at a particular time.

8. There is close agreement between teacher education toward the year 2000 as seen by the participants in this study and the goals of teacher education cited in one previous study.

The study was supported by a grant from the Alberta Advisory Committee on Educational Studies.

References

- Only those references directly pertinent to this study are cited. The general references most often referred to by the authors include Chapter XI, *Teacher Education in the Future*, (1975), Seventy-fourth Yearbook, N.S.S.E., and *The Journal Teacher of Education*, 1974, Vol. 25, No. 2, thematic section on *Futurism in Education*.
- Chalmers, R., & McPhie, G. *1,000,000 opinions: A report of the social values survey*. Montreal: The Protestant School Board of Greater Montreal, 1972.
- Clarke, S. C. T., & Coutts, H. T. The future of teacher education. *The Journal of Teacher Education*, 1971, 22 (4), Winter. Also reprinted in R. W. Hostrop (Ed.), *Foundations of futurology in education*. Homewood, Ill.: ETC Publications, 1973, 114-133.
- Coutts, H. T., & Clarke, S. C. T. *The goals of teacher education*. Research Monograph #20, The Alberta Teachers' Association. Edmonton, Barnett House, 1972.
- Coutts, H. T., Clarke, S. C. T., & Horowitz, M. The goals of education. Unpublished manuscript. The University of Alberta, 1974.
- Lauwerys, J. *The purpose of education: Results of a C.E.A. survey*. Toronto: Canadian Education Association, 1973.

R. O'REILLY

The University of Ottawa

Classroom Climate and Achievement in Secondary School Mathematics Classes

The purpose of this study is to show that the psychological climate of a classroom has an effect on the learning of students. Consequently teachers and supervisors should monitor climate and develop techniques for improving it. The subjects for the study were 1,100 secondary students in 48 mathematics classes in Eastern Ontario. Variables in the study included pupil scholastic aptitude, student achievement, classroom climate as measured by the Learning Environment Inventory (LEI) and pupil characteristics such as social background, liking for school and future orientation. (Dr. O'Reilly is Associate Professor in the Faculty of Education, Graduate Studies, at the University of Ottawa.)

The purpose of this paper is to examine classroom climate as an intervening, or process, variable in education. Administrators and supervisors examine pupil achievement when they assess educational programs. In their evaluation they also allow for the native ability and social background of students. Yet, although they claim to pay a great deal of attention to actual classroom behavior, they fail to examine in a systematic fashion the psychosocial environment of the classroom as a significant intervening variable. It is proposed here that climate is an important variable which should be monitored, manipulated and related to educational outcomes. In order to demonstrate this, data collected in an Ontario study will be utilized to examine the relationship between certain social and personal characteristics of pupils, the climate of grade 9 and 10 mathematics classes, and mathematics achievement of students.

Background

The most persistent questions faced by the educational researcher are those which ask, "What accounts for the variation in the learning of pupils achieved in the classroom?" Thus far, we have been quite successful in relating achievement to student aptitude and to such home factors as socioeconomic status, expectations for success and level of parents' education. Pupil aptitude and scores on I.Q. tests usually account for about 50 per cent in the variance in

school achievement whereas home factors account for 65 percent of the variance in standardized tests. It is on the basis of results such as these that social reformers state that schools do little to reduce effects of social inequalities. Jencks (1972) does suggest in his conclusions, that it is the social environment that immerses the school and the child which accounts for learning rather than curricula, teaching style or the availability of resources. He recommends that schooling be thought of as an end in itself rather than as a means to other ends such as cognitive development or as preparation for adult responsibilities. He goes on to say that although schools should have goals, for all enjoyable activities require purpose, looking back over his data, he concludes that it doesn't matter what those goals are.

Educators must study carefully the reports of studies such as Jencks; however, to accept his conclusions and the model of schooling implied in those conclusions would be to deny our profession as educators. Education is predicated on the premise that, as a result of performing given activities, pupils will change and the change will be in given direction. That is, schools aid children to develop in desired and desirable directions.

We do not deny the existence of heredity. But it is our point of departure; not our end-point.

We agree that the social context of learning is also important; but we deny that the macro-social environment is so overwhelming that the micro-environment of the school and of the classroom is insignificant for learning.

Modern behavioral science, since the time of Kurt Lewin, has been constructed on the dual premise that the social-psychological environment or climate of a work group strongly influences outcomes, and that this climate is influenced by the group leadership and is thus subject to manipulation (Lewin, 1947).

Although a number of methodologies have been devised to research these variables in schools, few studies have been designed to relate climate to learning outcomes. A recent survey located only 13 such studies in addition to a series to be discussed below. The thirteen studies indicated few clear-cut relationships between climate and achievement. The authors attribute this, among other reasons, to the poor validity of instruments, and the use of global concepts of climate rather than breaking it down into its components (Nielsen & Kirk, 1974). In another survey of the research, Rosenshine (1970) encountered similar difficulties.

One classroom climate rating form which has enjoyed some success is the *Learning Environment Inventory* (LEI), developed by Walberg and Anderson (1968) for the evaluation of the Harvard Project Physics. This instrument has been useful in relating climate to achievement; it has also been useful in that teaching methods also related to climate are measured by the LEI.

The final version of the scale contains 105 statements describing typical school classes. Each student expresses his agreement or disagreement on a four point scale. The items are grouped into 15 scales defined as follows:

1. *Cohesiveness*: The feeling of intimacy that has developed as a result of several individuals interacting over a period of time.
2. *Diversity*: The extent to which the class provides for a diversity of pupil interests and activities.

3. *Formality*: The extent to which behavior within the class is guided by formal rules.
4. *Speed*: The rate of progress of the class.
5. *Environment*: The physical environment, including the amount of space available and the type of recreational equipment.
6. *Friction*: The extent to which conflict may affect the behavior of the class.
7. *Goal Direction*: The recognition of goals and their subsequent acceptance by the group.
8. *Favouritism*: The extent to which pupils possess a low academic self-concept.
9. *Cliqueness*: Aims at revealing the extent to which cliqueness exists in a classroom and its influences on social interaction.
10. *Satisfaction*: The extent to which students like or dislike their class.
11. *Disorganization*: The extent to which students consider their class to be disorganized.
12. *Difficulty*: The relative perceived difficulty levels of various courses.
13. *Apathy*: Complements the cohesiveness scale, but also indicates if individuals within the class feel any affinity with class activities.
14. *Democratic*: Indicates the extent to which a "democratic" atmosphere exists within a classroom.
15. *Competitiveness*: The degree of competitiveness existing within the class.

Several reliabilities are available. For class groups, intraclass correlations range from .31 to .92 for each scale. Test re-test correlations for each scale range from .43 to .73. The LEI was successfully utilized in a variety of experimental and correlational studies, described by Anderson and Walberg (1974).

Design of the Study

The unit of analysis in this study is the classroom, and all scores are average classroom scores. Forty-eight classrooms from 12 secondary schools from four boards of education in Eastern Ontario constituted the sample. To fulfill certain requirements of the major project, a combination of random and judgmental sampling was employed. The forty-eight classes were equally divided between the grades 9 and 10 levels (or equivalents) and between advanced academically-oriented and general or terminal classes. A total of 1,100 students participated.

Data Collection

Data were collected from students with the following instruments:

1. *Stanford Achievement Test (SAT)*: Mathematics Form W.
2. Biographical data.
 - (a) parents' education
 - (b) dislike for school
 - (c) future planning
3. LEI.
4. *School and College Ability (SCAT) Test*, Series II.

Questionnaires and tests were administered in two sittings with each of the 48 classes. One sitting involved the administration of the *Stanford Achieve-*

ment Test—Mathematics—Part A, and the SCAT—Series II, Form 2A in Grade 10 Classes, Form 3A in Grade 9 Classes. The Stanford was administered to 2/3 of the class while the SCAT was administered to the remaining one-third. In the other sitting, the LEI and the Biographical Inventory were administered to one half of the class while other tests were administered to the remainder of the class. Males and females were proportionally represented in each segment of data collection.

The order of administering sets of instruments in the two settings was alternated to eliminate order effects. This method of collecting data was consistent with procedures used by Walberg and Anderson (1968), and was economical in terms of research effort.

Analysis of Data

In order to analyze the results obtained in this project, the framework of INPUT-PROCESS-OUTPUT will be used.

Input variables. Input variables are usually beyond the control of the educator. Those given are: (1) scholastic aptitude (I.Q.) and (2) biographical indices. Correlations with Achievement are shown in Table 1.

TABLE 1
MATRIX OF INPUT AND OUTPUT VARIABLES

Variable	2	3	4	5	R
1. Achievement	.83*	.48*	-.41*	-.26	.88*
2. Scholastic Aptitude	1.00	.45*	-.33*	-.26	
3. Parents' Education		1.00	-.08	-.30*	
4. Dislike for Schooling			1.00	.23	
5. Career Planning				1.00	

* $p < .05$

N = 48

Scholastic aptitude is related to achievement (verbal, $r = .62$; mathematical, $r = .86$). Of the biographical indices, parents' education is related to achievement ($r = .48$); dislike for schooling is similarly related ($r = -.41$). The third biographical variable, labeled "career planning", was intended to obtain an index of the student's future orientation, but it proved to be unsatisfactory in this study.

Taken altogether, these input variables account for 77% of the variance in achievement ($R = .88$).

Process Variables. The process variables are the fifteen scales of the LEI. Taken together, they account for 67 per cent of the variance in achievement ($R = .82$).

The LEI variables are not independent of input variables. This one might expect, as pupils bring a great deal of their world into the classroom. Pupils have a significant effect on the climate of the classroom.

Eleven of the 15 LEI variables are associated with input variables (variables 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 and 15). Scholastic aptitude is associated with

nine LEI scales. All input variables are related to the variables Friction, Satisfaction and Disorganization. In addition, “Dislike for schooling” is related to Favoritism, Apathy and Democratic variables. “Parents’ education” is related to an additional three variables: Goal Direction, Difficulty and Apathy.

TABLE 2
CORRELATIONS OF INPUT-PROCESS AND OUTPUT VARIABLES

LEI Scales	IQ	BI(1)	BI(2)	BI(3)	Achievement
1. Cohesiveness					.30
2. Diversity					
3. Formality					
4. Speed					
5. Environment			-.46		.32
6. Friction	-.54	-.31	.45	.35	-.59
7. Goal Direction		.30			
8. Favoritism	-.42		.45		-.44
9. Cliqueness	-.33				-.34
10. Satisfaction	.43	.31	-.49	-.39	.45
11. Disorganization	-.50	-.35	.48	.48	-.52
12. Difficulty	.40	.31			.33
13. Apathy	-.52	-.40	.42		-.44
14. Democratic	.33		-.41		.47
15. Competitiveness	.35				
R					.82

IQ is SCAT-Total
BI(1) is Parents' Education
BI(2) is Dislike for Schooling
BI(3) is Career Plans.

Only significant correlations are shown. N = 48

Ten LEI variables are related to achievement. Achievement in mathematics classes is characterized by high scores on the following variables: Cohesiveness (.30); Environment (.32); Satisfaction (.45); Difficulty (.33); and Democratic (.47). High achieving classes are noted by low scores on Friction (-.59); Favoritism (-.44); Cliqueness (-.34); Disorganization (-.52); and Apathy (-.44).

Discussion

There are a number of concerns here for the administrator.

If Input is related to both Process and Achievement, can the teacher be in any way held accountable for the classroom climate or for the achievement of pupils?

In order to answer that question, three elements must be considered. First, do input variables account for all the variance? Second, can teachers modify the environment of the classroom? Third, can school be a fruitful experience for all pupils?

One technique available to answer the first question is to compute partial correlations, a technique which first subtracts the common variance which two

variables share with a third variable. Walberg and Anderson (1972) computed correlations between LEI variables and year-end provincial examination scores covering eight subject areas with intelligence partialled out of the relationship. Their sample was 64 grade 10 and 11 classes from eight schools in Montreal. Twelve of the LEI scales were significantly related to overall achievement. The size of the correlations, on the average, were 10% smaller than the correlations between achievement and LEI scales. The data from the O'Reilly and Garland (1974) is less satisfying (Table 3). Correlations between achievement and LEI scales with total SCAT scores partialled out yielded only three significant correlations (Diversity, $-.35$; Friction, $-.30$; Democratic, $.36$). In this case, the Walberg and Anderson data should be given greater credence, since they related their scores to actual examination results, whereas the O'Reilly and Garland study utilized standardized achievement tests which are somewhat confounded with scholastic aptitude tests. The correlation between achievement and aptitude in the O'Reilly and Garland study was $.82$ and in the Walberg and Anderson study, it was only $.38$.

TABLE 3
CORRELATIONS BETWEEN LEARNING ENVIRONMENT SCALES AND
MATHEMATICS ACHIEVEMENT SCORES (SAT) AND PARTIAL COR-
RELATIONS

Environment Scales	Mathematics Achievement	Partial Correlations, Controlling for...			
		Intelligence SCAT-Total	Parents' Education	Dislike for Schooling	Future Plan Orientation
1. Cohesiveness	.30		.29	.25	.30
2. Diversity		-.35			-.26
3. Formality					
4. Speed					
5. Environment	.32		.32		.28
6. Friction	-.59	-.30	-.53	-.49	-.55
7. Goal Direction					
8. Favoritism	-.44		-.38	-.32	-.41
9. Cliquesness	-.34		-.27		-.32
10. Satisfaction	.45		.36	.31	.39
11. Disorganization	-.52		-.43	-.40	-.47
12. Difficulty	.33			.40	.31
13. Apathy	-.44		-.31	-.32	-.41
14. Democratic	.47	.36	.41	.36	.43
15. Competitiveness	.25			.31	
R	.82				

$r_{.05} = .29$ Blanks indicate $r < .25$. $N = 48$

In considering correlations between achievement and aptitude when biographical indices are partialled out, the effects of the LEI scales remain pronounced. When controlling for parents' education, seven of the ten LEI scales remained significantly correlated. Overall the correlations are approximately 10% lower.

When controlling for the pupils' dislike for schooling, eight correlations remain significant.

Controlling for future plan orientation produces almost no changes.

Thus, although students' characteristics do predict class environment to some extent, there is still room for the effect of teacher and school controlled variables upon academic achievement.

The answer to the second question cannot be answered by the data presented here, but there is the inference that teachers can influence climate and consequently achievement. Further, it is axiomatic in the literature on leadership, supervision and management, that leaders do have an impact on both group climate factors and on achievement (Likert, 1967). Early studies by Anderson and Brewer (1945) established that a teacher can have an effect on classroom climate very early in the school year. It would seem logical to assume that the climate of a classroom is the result of the interaction of pupil and teacher characteristics and that the influence of the teacher would not be insignificant. Consequently a study of classroom climate early in the school year would be advisable.

To the third question, "Can the school be a fruitful experience for all students?" the reply must hinge on the level of ability of the teacher. In an age when teachers were less well trained and when there were few opportunities for professional development, it could have been expected that teachers would easily teach the able and motivated student, whereas the others, subjected to repeated failure, would eventually drop out. Today it is difficult to excuse a university trained profession with ample opportunities for professional development, for failure to provide the majority of all pupils with satisfying learning experiences within the school.

The identification of important aspects of climate gives an added tool to teachers and supervisors. In management theory, it has been shown that trends in the change of direction in climate can be detected fairly early, whereas the effects of such changes on productivity might be delayed by as much as two years (Likert, 1967). Educators could begin to systematically monitor classroom climate and attempt to influence its nature. Expertise to do this is rapidly appearing in the education profession (Phi Delta Kappa, 1974). A supervisory program could be designed which would call for early monitoring of classroom climates, and if the situation warrants it, expert assistance, perhaps in the form of a school guidance teacher or an organizational development professional, would assist in a program of climate change.

To return to the discussion of the study by Jencks, on the basis of the material presented in this paper, we can conclude:

1. Classroom climate is a significant factor in pupil achievement.
2. Although pupil personal and social characteristics are important correlates of achievement, climate in some instances is a more important factor.
3. Although input, process and output are all intercorrelated, process variables have an independent effect on achievement.
4. Education includes psycho-social factors as well as academic factors and the two are related.

References

- Anderson, G., & Walberg, H. J. Learning environments. In H. J. Walberg (Ed.), *Evaluating educational performance*. Berkeley: McCutchan Publishing, 1974, 81-98.
- Anderson, H. H., & Brewer, J. E. Studies of teacher classroom personalities. *Applied Psychology Monograph*, 1945, 6.
- Jencks, C., et al. *Inequality: A reassessment of the effect of family and schooling in America*. New York: Basic Books, 1972.
- Lewin, K. Frontiers in group dynamics. *Human Relations*, 1947, 1, 5-41.
- Likert, R. *The human organization*. Toronto: McGraw-Hill, 1967.
- Nielsen, H. D., & Kirk, D. H. Classroom climates. In H. J. Walberg (Ed.), *Evaluating educational performance*. Berkeley: McCutchan Publishing, 1974, 57-79.
- O'Reilly, R., & Garland, P. *Paradigm for evaluation in the high school*. Ottawa: University of Ottawa, 1974.
- Phi Delta Kappa. *School climate improvement*. Bloomington, Ind.: Phi Delta Kappa, 1974.
- Rosenshine, B. Evaluation of instruction. *Review of Educational Research*, 1970, 40, 279-300.
- Walberg, H. J., & Anderson, G. Classroom climate and individual learning. *Journal of Educational Psychology*, 1968, 59, 414-419.
- Walberg, H. J., & Anderson, G. Properties of the achieving urban classes. *Journal of Educational Psychology*, 1972, 61, 381-385.

D. F. ROBITAILLE

University of British Columbia

Classroom Personality Patterns of Teachers of Secondary Mathematics

The 167 teachers of secondary mathematics in a public school system in Montreal, Quebec were rated by four judges working independently of one another with usable data being obtained on 127 of the teachers. Each teacher was assigned the sum of his four ratings as an effectiveness score. The thirty teachers with the lowest effectiveness scores were designated as minimally effective and the twenty-five with the highest scores as highly effective.

The secondary school principals judged the teachers' classroom personality patterns by assigning each teacher to one of six personality categories. Three of these categories had positive denotations and three had negative denotations. Results showed that all twenty-five effective teachers were assigned positive categories with over 80% being rated as responsible, businesslike, and systematic in their teaching. Over 60% of the ineffective teachers were assigned to negative categories with most of them being rated as dull and routine teachers. (Dr. Robitaille is Assistant Professor in the Faculty of Education, The University of British Columbia.)

The area of teacher effectiveness has produced probably more research than any other in the field of education. In spite of all this effort, until quite recently writers in the area have been of the opinion that very little was known for certain about the ingredients constituting effective teaching or about the relationship between teacher personality and teaching effectiveness (Getzels & Jackson, 1963).

It may be the case that one of the reasons for the lack of compelling findings from teacher effectiveness research has been that the research questions have frequently been too broad. Much previous research seems to be based on the assumption "... that there is some magic variable that applies to all of teaching, for all students, at all grade levels, in all subject matters, and for all objectives" (Gage, 1972).

It is probably not possible to define and characterize teaching effectiveness in general, at least until a series of preliminary questions has

been answered. Some such preliminary questions might address specific aspects of the teaching act. Gage, for example, has discussed the use of "micro-criteria".

Rather than seek criteria for the overall effectiveness of teachers in the many, varied facets of their roles, we may have better success with criteria of effectiveness in small, specifically defined aspects of the role. (Gage, 1972, p. 95)

These aspects of the teaching act might include such "technical skills" as explaining, questioning, obtaining student participation, providing feedback, and so on.

A second set of preliminary research questions, not necessarily completely disjoint from the first, might address the area of grade level and subject matter differences in teaching. Thus, it may be fruitful to attempt to characterize the effective teaching of secondary mathematics, for example, before attempting to describe effective teaching in general. In other words, it may be the case that the characteristics of effective teachers of a given discipline at a given grade level are quite different in many ways from those of effective teachers of the same subject at a different grade level or from those of effective teachers of a different discipline at any grade level.

There is some evidence to show that teachers of a given subject at a given grade level do exhibit personality characteristics which differ from those of other teachers at that grade level in other subject areas. In the Teacher Characteristics Study (Ryans, 1960), three major patterns of teacher classroom behavior were identified. These were:

TCS Pattern X: warm, understanding, friendly versus aloof, egocentric, restricted teacher classroom behavior;

TCS Pattern Y: responsible, businesslike, systematic versus evading, unplanned, slipshod teacher classroom behavior;

TCS Pattern Z: stimulating, imaginative versus dull, routine teacher classroom behavior.

Ryans' results indicate that "superior" teachers of secondary mathematics and science score significantly higher than "poor" teachers of these same disciplines on all three of the TCS patterns: X, Y, and Z. The teachers who participated in this aspect of the Teacher Characteristics Study were nominated by their respective principals as being outstandingly superior or poor (Ryans, 1964).

The present paper is a report of an attempt to discover which one of the TCS patterns best describes the classroom behavior of a highly effective teacher of secondary mathematics, and which one best describes the classroom behavior of a minimally effective teacher of secondary mathematics. Data such as these may be useful in determining some of the parameters governing teacher effectiveness in the area of secondary school mathematics.

Method

All one hundred sixty-seven teachers of secondary school mathematics, grades 8-12, in a public school system in Montreal, Quebec were rated by each of four judges working independently of each other. The raters were the Assistant Director of Secondary Education, the two Mathematics Coordinators for the system (including the author), and the principal of the

school where the teacher being evaluated taught. Thus each teacher was evaluated by his or her own principal as well as by three persons who had system-wide responsibilities.

The evaluators rated each teacher on a five-point scale ranging from 1 to 5. A score of 1 indicated minimal teaching effectiveness, while a score of 5 indicated a very high degree of teaching effectiveness. No criteria of teaching effectiveness were provided for the raters. They were merely asked to provide their personal opinions of each teacher's effectiveness by assigning a score of between 1 and 5. When the ratings had been completed, each teacher was assigned the sum of his four ratings as an effectiveness score. The lowest possible effectiveness score was 4 and the maximum was 20. Of the total of 167 teachers, usable data were obtained on 127; that is, approximately 76% of the total. The remaining 40 teachers were new to the system and, in each case, at least two of the raters felt unqualified to rate them.

In a related study (Robitaille, 1975), those teachers with effectiveness scores in the range 16-20 were designated as highly effective teachers and those whose effectiveness scores fell in the range 4-9 were designated as minimally effective. Analysis of background data and the results of paper and pencil tests administered to these teachers strongly supported the evaluations given by the raters. That is, the effective teachers were significantly better than the minimally effective teachers on several measures. The testing instruments used for this purpose included the Contemporary Mathematics Test (Massie, 1967), the Teaching Situation Reaction Test (Duncan & Hough, 1966), the Minnesota Teacher Attitude Inventory, and the Purdue Teacher Opinionnaire. Similarly, data obtained from classroom observations of these same teachers produced significant differences in favor of the teachers rated as highly effective with respect to the extent of student participation in their classes.

At the same time as the principals rated each of their teachers, they were asked to select from a given list the personality pattern which best described that teacher's classroom behavior. For this purpose, each of the three TCS patterns was divided into two parts, one with a positive denotation and the other with a negative one. The six resulting categories were: XA—warm, friendly, understanding teacher behavior; XB—aloof, egocentric, restricted teacher behavior; YA—responsible, businesslike, systematic teacher behavior; YB—evading, unplanned, slipshod teacher behavior; ZA—stimulating, imaginative teacher behavior; ZB—dull, routine teacher behavior.

Results and Discussion

The distribution of effectiveness scores is shown in Table 1. Thirty teachers out of the total pool of 127 were rated as minimally effective; that is, their effectiveness scores fell within the range 4-9. Twenty-five teachers were judged to be highly effective; that is, their effectiveness scores fell within the range 16-20. The mean effectiveness score was 12.5 and the standard deviation of the scores was 3.6.

A chi-square test for goodness of fit supported the hypothesis that the distribution of effectiveness scores was not significantly different from a

TABLE 1
DISTRIBUTION OF EFFECTIVENESS SCORES

Score	Frequency	Score	Frequency
4	1	13	14
5	3	14	9
6	2	15	17
7	6	16	6
8	7	17	10
9	11	18	2
10	4	19	5
11	12	20	2
12	16		

N = 127

normal distribution. In applying this test, the three lowest categories (4, 5, and 6) were grouped together because of the small number of cases involved. Categories 18, 19, and 20 were also grouped together for the same reason.

A coefficient of inter-observer reliability was computed using a technique suggested by Kerlinger (1964). This involves the use of a two-way analysis of variance (see Table 2) in order to compute the required value.

$$r = 1 - \left(\frac{V_e}{V_{ind}} \right) = 1 - (0.43/6.56)$$

V_e indicated "the estimate of the error variance" and V_{ind} denotes "the variance resulting from individual differences." This technique produced a 0.94 coefficient of inter-observer reliability indicating a high degree of agreement among the four evaluators.

TABLE 2
ANALYSIS OF VARIANCE OF EFFECTIVENESS SCORES

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Squares
Teachers	126	407.92	3.24
Effectiveness Scores	3	19.69	6.56
Residual	378	161.81	0.43

The findings regarding the teacher's classroom personality patterns as judged by their principals are summarized in Table 3.

The twenty-five teachers who formed the pool of highly effective teachers were all rated by their principals as belonging to one of the three positive categories XA, YA, and ZA. Of the 25, 80% were judged to exhibit responsible, businesslike, and systematic behaviour (category YA) as the predominant trait in their teaching.

Of the thirty teachers judged to be minimally effective over 60%, 19 out of 30, were seen by their principals as exhibiting one of the three negative

TABLE 3
MATRIX OF EFFECTIVENESS SCORES VERSUS PERSONALITY PATTERNS

Personality Patterns	Low (4-9)	Moderate (10-15)	High (16-20)	
XA	10	25	6	41
XB	4	3	0	7
YA	1	30	16	47
YB	4	0	0	4
ZA	0	5	3	8
ZB	11	9	0	20
	30	72	25	

personality patterns in their teaching. More than half of these, 11 out of 19, were seen as being dull and routine in their teaching. Of the eleven minimally effective teachers who were awarded a positive personality pattern over 90%, 10 out of 11, were judged to be warm, friendly and understanding in their teaching. It is possible that some of these teachers were judged to have an excess of personality pattern XA and for this reason were given low effectiveness scores. Also noteworthy is the fact that only 6 of the 25 most effective teachers were placed in category XA.

Two categories, namely YA (responsible, businesslike, and systematic) and ZB (dull and routine), together include more than 50% of the teachers, 67 out of 127 to be precise. It may be that the distinction between these two categories is not as clear as it could be and that some principals had difficulty differentiating between the two. In other words, an evaluator might find it difficult to differentiate between a teacher whose teaching is "routine" and one whose teaching is "systematic" or "businesslike". If it is indeed the case that there is an ambiguity here and that the distinction between the two categories is vague, then these principals may be saying that the majority of these teachers of secondary mathematics exhibit classroom behavior which could be characterized as dull and routine.

Of the three positive categories, ZA (stimulating and imaginative teacher behavior) was utilized least. Only eight teachers, approximately 6% of the total, were seen as being predominantly stimulating and imaginative in their teaching behavior. More interestingly, of these eight only three were rated among the highly effective teachers. This finding seems to indicate that, in the opinion of principals, very few teachers of mathematics are predominantly stimulating and imaginative in their teaching and also that stimulating and imaginative teachers are frequently *not* among those judged to be the most effective teachers.

From another point of view, the finding is that only 3 of the 25 most effective teachers are judged to be stimulating and imaginative. This does not necessarily mean that the other 22 highly effective teachers are not stimulating and imaginative, but rather that one of the other two personality patterns XA and YA is predominant in their teaching behavior.

Still, the result does seem to say that principals find very few of their teachers of mathematics to be particularly stimulating and imaginative in their teaching. Such a finding should be of interest to those who are involved in the training of prospective teachers of secondary mathematics as well as to those involved in in-service programs for teachers.

Summary

The overwhelming majority of the most effective teachers were judged by their principals to be responsible, businesslike, and systematic in their teaching of mathematics. The other two positive categories, XA and ZA, included only 20% of the highly effective teachers.

The results for the minimally effective teachers are not as clear cut. One third of these teachers were rated as warm, friendly and understanding, a positive category, while another 11 of the group of 30 were seen as exhibiting dull and routine teaching behavior, category ZB.

References

- Biddle, B. J., & Ellena, W. J. (Eds.) *Contemporary research on teacher effectiveness*. New York: Holt, Rinehart and Winston, 1964.
- Duncan, J. K., & Hough, J. B. Technical review of the teaching situation reaction test. Unpublished manuscript, The Ohio State University, 1966.
- Gage, N. L. *Teacher effectiveness and teacher education: The search for a scientific basis*. Palo Alto, Calif.: Pacific Books, 1972.
- Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), *Handbook of research on teaching*. Chicago: Rand McNally, 1963.
- Kerlinger, F. N. *Foundations of behavioral research*. New York: Holt, Rinehart and Winston, 1964.
- Massie, R. O. The construction and use of a test to evaluate teacher preparation in modern mathematics. Unpublished doctoral dissertation, University of Nebraska, 1967.
- Robitaille, D. F. Criteria for assessing the effectiveness of teachers of secondary mathematics. *Journal for Research in Mathematics Education*. 1975, 6, 77-87.
- Ryans, D. G. *Characteristics of teachers*. Washington: American Council on Education, 1960.
- Ryans, D. G. Research on teacher behavior in the context of the teacher characteristics study. In B. J. Biddle & W. J. Ellena (Eds.), *Contemporary research on teacher effectiveness*. New York: Holt, Rinehart and Winston, 1964.

J. P. KLEIN

University of Toronto

Socratic Dialogue vs. Behavioural Practice in the Development of Coping Skills

An attempt was made to increase the competency skills of 44 male students in grades 7 to 9 through the Socratic teaching of the theory of security and independence, the behavioural practice of coping skills, or both. At the end of six half-hour sessions no significant improvement was found in math grades, but directional and significant differences demonstrated that the groups receiving dialectic training increased confidence in their problem solving capacities, and showed a heightened sense of security and self-directedness. (Dr. Klein is with the Institute of Child Study, Faculty of Education at the University of Toronto.)

With the increasingly widespread recognition that the results of traditional forms of psychotherapy with children are a serious disappointment to our optimistic imagination (Levitt, 1963), some clinicians emphasize the need for experimental modes of treatment (Lazarus, 1971) while others shift the entire focus from remediation to prevention (Kohlberg, 1972). Accompanying the latter thrust of inquiry is the substitution of cognition for affect as the major variable of interest. Disruptive emotions are viewed as products of irrational thinking (Ellis, 1962) or maladaptive coping responses (Lazarus, 1966). The road to health is the one which secures ego strength; the child possessing those adaptive skills which lead to the mastery of life stress situations will be free from psychological problems. He has ego strength. Those possessing these skills in smaller degrees are increasingly inclined to utilize inadequate acts of mastery called defenses; they employ no strategies which lead them on to depression; or they act with strategies curiously inappropriate to the stress event called psychosis.

While many investigators explore the life-history antecedents and personality correlates of coping skills, to the author's knowledge this is the first study that attempts to test the hypothesis that ego strength can be experimentally developed.

The first method chosen is the Socratic dialogue. While the juxtaposition of seemingly contradictory arguments in order to stimulate debate leading towards a truthful conclusion has a long classroom history, the method and its purpose as stated by Socrates (Plato, 1927) have never been accurately followed. Socrates never taught before he exhorted and provoked his student into a realization of his ignorance about something of vital importance to his life. Modern psychology has elaborated on the extent to which ego-involvement raises the attractiveness of a task (Ausubel & Sullivan, 1970). In addition, Socrates never used the techniques as a means of communicating factual information. Education was a way of bringing out qualities of character already possessed but that lay dormant and undifferentiated within a person. Socratic dialogue was a way of becoming somebody, not knowing something.

The second method is a behavioural approach which presents a good contrast to much of the theoretical thinking about ego strength. By getting a student to perform directly the final mastery behaviours, we shift from an analysis of the intervening variables that modulate coping skills as they are explained by the ego psychologists (White, 1974), to a simple habitual exercise of those terminal acts which constitute adaptation to life problems.

Method

Subjects

All male students in grades 7 to 9 of a Toronto junior high school were told that for their voluntary co-operation in an experimental project, they would receive their choice of a gift worth five to ten dollars. The mean age of the 52 volunteers was 13.0 years. The students were ordered alphabetically by last name. Blocks of four students were then taken, and a table of random numbers was used to assign the students within blocks to the experimental treatments. Each of the four groups had 13 Ss. The author served as *E*.

Design and Procedure

The four treatment conditions were (a) Socratic learning of the Theory of Security, (b) behavioural training in coping skills, (c) equal amounts of training in (a) and (b), (d) a no treatment control group. Each treatment group met every third school day for 30 minutes during school hours over the span of a month. This resulted in a total of six half-hour sessions for each treatment group. Two one-hour testing sessions were held for all Ss: one directly prior and one directly following the treatment program.

Socratic Dialogue

The Institute of Child Study Security Test (Grapko, 1957) called "The Story of Jimmy" is a written projective measure in which the *S* ranks five possible solutions to a given problem in the life of Jimmy. The alternatives reflect varying degrees of maturity in the manner of accepting consequences for one's own decisions. Those six problem situations most closely related to school life were chosen (numbers 1, 4, 5, 8, 9, 10). For each session the Ss read one of the problems to themselves and ranked the given choices from the one most to the one least likely for Jimmy to choose. The

E, following the Socratic method as outlined in Oliver and Shaver (1966), then called upon an *S* to state his lowest ranked choice and give the reason for his selection. *E*, assuming the Socratic adversarial role, questioned the *S* by presenting a parallel but slightly different situation to which the *S* would have responded differently. The *S* was then asked to state the principle which would lead him to a different decision in each of the two parallel cases. Other *Ss* often intervened to present their different choice for lowest rank, or to criticize the principle enunciated by the first *S*. *E*, interjecting statements about the seriousness of being correct in matters of personal responsibility and the importance of acting like an adult, always concluded by teaching the principle which security theory attributed to that rank. Then the group went on to the second from last choice and proceeded in the above pattern until all five choices were discussed in the half hour session. The six situations thus occupied the six treatment sessions.

Behavioural Training

The *E* requested the *Ss* to list those aspects of academic life for which they were generally held responsible. *E* added those omitted by *Ss* so that the list included controlling academically irrelevant behaviour, paying attention in class, raising your hand to ask questions when unclear about the lesson being taught, doing the homework regularly and independently, and studying before an exam. *Ss* volunteered responsibilities often evaded at home such as straightening the bedroom, dumping the garbage or the brief caring for a pre-school sibling. To insure confidentiality and honesty at the beginning of a session, each *S* would write on a sheet of paper his name and those responsibilities which he had performed since the last session. The *E* requested *Ss* to report verbally any difficulties they had in fulfilling these responsibilities and suggested ways of overcoming them. *E* suggested the four times per week math class as a focus for examining barriers to achieving the behavioural goals.

Combined Condition

Ss in this group spent the first half of each session at the Story of Jimmy and the second half at behavioural training. They covered the same ground as the preceding groups, but learned it less intensively.

Dependent Variables

With the exception of math grade, the following measures were administered to all *Ss* directly before and once again directly after the treatment program.

Anxiety Scale. The Children's Manifest Anxiety Scale (CMAS) (Casteneda, McCandless & Palermo, 1956) was used as an overall measure of anxiety and in security.

Problem Solving. A Child Attitude Inventory for Problem Solving (CAPS) developed by R. S. Crutchfield and M. L. Covington (Johnson & Bommarito, 1971) is a 33 item questionnaire measuring beliefs about problem solving in general and one's own ability at problem solving in particular.

Responsibility for Achievement. The Intellectual Achievement Responsibility Questionnaire (IAR) (Crandall, Katkovsky, & Crandall, 1965) is a

34 item questionnaire measuring the extent to which a child attributes his intellectual progress to himself or to external factors beyond his control. Seven scores can be derived by dividing the items into those mentioning behavioural effort (e) and those not mentioning a specific cause (u for undifferentiated), and events with positive (+) or negative (-) outcome. Thus I+e is arrived at by adding the S's internal responses (I) to items in which the positive events (+) stem from an internal response indicating behavioural effort (e). Similarly, scores for I+u, and I-u are calculated. The two I+ subsets can be added to give an additional score as can the two I- subsets; and an overall score can be obtained by adding these two. Every scored item raises the S's measure of inner-directedness.

Math Grade. By prior arrangement with the math teacher but unknown to the Ss, a new unit of math study began with the initiation and ended at the conclusion of the treatment program. Marks in math on the Fall report card were the pretreatment measure of academic achievement, and an objective quiz on the unit, administered and scored by the teacher, served as the post-treatment measure.

Results

Three Ss in Group 1, four Ss in Group 3, and one S in Group 4 were excluded from the study because they either completed the tests incorrectly or missed more than two treatment sessions. For the remaining 44 Ss, an analysis of covariance was performed using the pretreatment measure as covariate and post-treatment measure as criterion for each of the dependent variables. All tests were nonsignificant with the following exceptions: a directional effect was evident on the CMAS, on the CAPS, and on the Iu+ (see Table 1). The adjusted criterion group means are presented in Table 2.

TABLE 1
ANALYSIS OF COVARIANCE

Source	Sums of Squares	df/df	F
Children's Manifest Anxiety	139.42	3/39	2.29*
Attitude for Problem Solving	815.07	3/39	2.31*
IAR (Iu+)	12.74	3/39	2.56*

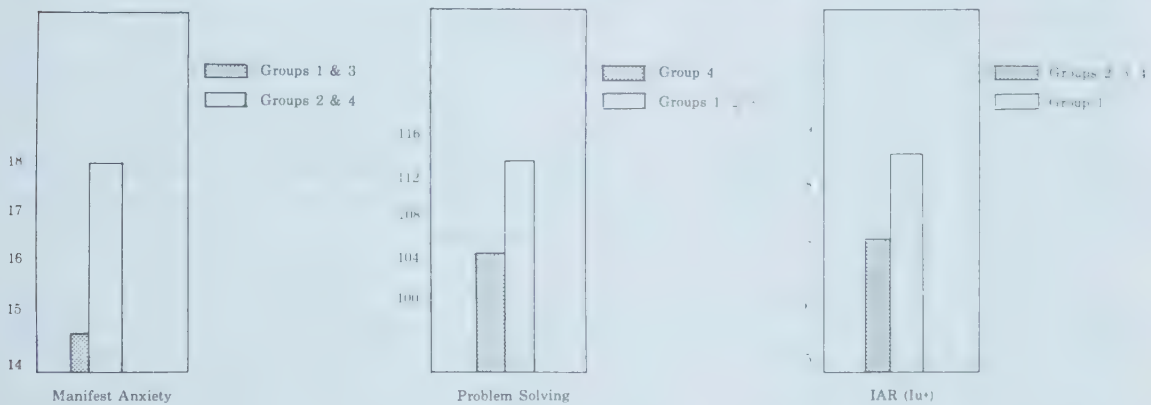
* $p < .10$

TABLE 2
ADJUSTED CRITERION GROUP MEANS

	Group 1	Group 2	Group 3	Group 4
Children's Manifest Anxiety	14.65	18.58	14.44	17.66
Attitude for Problem Solving	115.96	110.95	112.40	104.02
IAR (Iu+)	8.49	7.07	6.98	7.35

A Duncan's Multiple Range test was performed for each of these three variables to identify the source of differences between groups. On the CMAS, Groups 1 and 3 formed a homogenous subset as did Groups 2 and 4. The former subset showed reduced anxiety. The subsets differed directionally ($p < .10$) at their closest points (Groups 1 & 4), and significantly ($p < .05$) at their most distant points (Groups 2 & 3). An analysis of the CAPS revealed one homogenous subset: Groups 1, 2, & 3 not differing amongst themselves, but Group 4, scoring the lowest for problem solving attitude differing significantly ($p < .05$) from Group 1 and directionally ($p < .10$) from Group 2 and from Group 3. An analysis of the IAR (Iu+) showed one homogenous subset: Groups 2, 3, and 4 not differing amongst themselves, but Group 1, showing the highest rating for academic responsibility, differing significantly ($p < .05$) from Group 2, 3, and 4. These effects are presented in Figure 1. Group means of homogenous subsets are averaged for easy identification.

FIGURE 1
A COMPARISON OF GROUP MEANS FOR THREE DEPENDENT VARIABLES



A Pearson Product-Moment correlation was computed between the pretests of all the dependent variables and the pretest to post test difference scores for these variables. Only those correlations relevant to the significant dependent variables and having a probability level of at least .05 are reported. In general these moderate correlations ranged from .30 to .50. Anxiety was negatively correlated with IAR positive (+) scores and with CAPS scores. Iu+ was positively correlated with other IAR scores and with CAPS scores. With reference to change scores, lesser anxiety accompanied increased scores on the CAPS. Increments in Iu+ were accompanied by increments in CAPS scores. No math measures correlated significantly with any other score.

Of the written reports for the performance of behavioural tasks of responsibility handed in by Groups 2 + 3, over 90% claimed successful completion of the tasks. But the language employed by the Ss so closely resembled the phrases used during the treatment session that their validity is suspect. The Ss' verbal reports about problems encountered in the completion of given tasks indicated a genuine involvement with the desired goals.

Discussion

Emergent from the pattern of results is the conclusion that the Socratic dialectic study of the "Story of Jimmy" is the most potent element in building coping skills. Its intensive study in Group I provides the first evidence that locus of reinforcement can be experimentally altered towards greater self-reliance. Its efficiency in strengthening security (reducing anxiety) appears in groups receiving smaller (Group 3) and greater (Group 1) amounts of study. Along with behavioural practice, such study also increases the self estimate of problem solving capacity and decreases the subjective estimate of the difficulty in solving problems in general.

If adaptation is accepted as the global concept under which various coping and mastery skills can be subsumed, a taxonomy of ego strength may be expounded. Following Piaget's developmental notion of adaptation as a form of compromise between the individual and his environment, a minimal condition necessary for successful adaption is the accurate understanding of the nature of the person-environment conflict which has produced the stress. It is hard to imagine a child making an adept adjustment to conflict if the parameters are unspecified or unclear to him. Secondly, the resolution of the stress event often assumes that the individual must at times relinquish self-direction in favour of being controlled by some external force. At this point the assimilation-accommodation equilibrium is upset by the predominance of the latter. Optimally speaking, adaptation must include a compromise in which the individual maintains a balance between freedom of action and compliance as expressed in the ability to use help.

In this light it is easy to see why the "Story of Jimmy" is so excellent a device for building (as well as measuring) adaptation. The various vignettes specify a conflict situation between the child and his environment. Each of the alternate solutions offered reflect one of the five levels of independence that graduate from independent security (assimilation of the environment to current cognitive schema) through mature dependent security (sharing the consequences of decisions) to insecurity (surrender of autonomy.) Far from emitting a negative reaction to being pushed beyond their present level of security, the students relished the understanding of social requirements for mature adult behavior. Maybe Socrates was right in assuming the existence of ego strength which only needed cultivation. This cognitive growth justifies the admission of intervening variables as a necessary component of any short-term program for enhancing adaptability to stress. Behavioural practice may develop such intellectual strength (all treatment groups showed gains on CAPS) but only in a retrospective trial and error attempt to formulate guiding principles. Security theory presents them at the outset.

How can we conclude without paying our respects to the situational complexities of human behaviour? Lazarus (1974) claims that one needs to gauge the entire situation before evaluating the level of coping efficiency. A common response to stress is an increase in affiliative behaviour. The students in the study recognized this difficulty when they claimed that less mature forms of security such as begging to be forgiven were generally more effective with teachers. Schools preach and practise the ethic that the meek shall inherit the earth, a maladaptive gesture in the business world.

Situational variance may account for the negative results with math grades. Perhaps previous knowledge of math is a stronger determinant of progress than any coping device. Future work should attempt to devise sound behavioural measures of adaptability. Lazarus (1974) has already warned us away from global measures, having discovered that the parts of an overall coping behaviour may serve antithetical functions. Only longitudinal data will ascertain whether gains in adaptability are permanent, and whether Kohlberg (1972) is correct in hypothesizing that the presence of various forms of competence is the best predictor for the absence of mental illness in adults.

This study was supported by the Ontario Educational Research Council.

References

- Ausubel, D. P., & Sullivan, E. V. *Theory and problems of child development*. New York: Grune & Stratton, 1970.
- Castaneda, A., McCandless, B. R., & Palermo, D. C. The children's form of the manifest anxiety scale. *Child Development*, 1956, 27, 317-326.
- Crandall, V. C., Katkovsky, W., & Crandall, V. J. Children's beliefs in their own control of reinforcement in intellectual-academic achievement situations. *Child Development*, 1965, 36, 91-109.
- Ellis, A. *Reason and emotion in psychotherapy*. New York: Lyle Stuart, 1962.
- Grapko, M. F. The Institute of Child Study Security Test, Elementary Form—*The story of Jimmy; Manual*. Toronto: University of Toronto Press, 1957.
- Johnson, O. G., & Bommarito, J. W. *Tests and measurements in child development: A handbook*. San Francisco: Jessey-Bass, 1971.
- Kohlberg, L. The predictability of adult mental health from childhood behavior. In B. B. Wolman (Ed.), *Manual of child psychopathology*. New York: McGraw-Hill, 1972.
- Lazarus, A. A. *Behaviour therapy and beyond*. New York: McGraw-Hill, 1971.
- Lazarus, R. S. *Psychological stress and the coping process*. New York: McGraw-Hill, 1966.
- Lazarus, R. S., Averill, J. R., & Opton, E. M., Jr. The psychology of coping: Issues of research and assessment. In G. V. Coelho (Ed.), *Coping and adaptation*. New York: Basic Books, 1974.
- Levitt, E. E. Psychotherapy with children: A further evaluation. *Behaviour Research & Therapy*, 1963, 1, 45-51.
- Oliver, D. W., & Shaver, J. P. *Teaching public issues in the high school*. Boston: Houghton Mifflin, Co., 1966.
- Plato, *Selections*. R. Demos (Ed.) New York: Scribner's, 1927.
- White, R. Strategies of adaptation: An attempt at systematic description. In G. V. Coelho (Ed.), *Coping and adaptation*. New York: Basic Books, 1974.

E. D. HOBBS

University of British Columbia

Methodological Problems in Conservation Testing With Particular Reference to Volume Conservation

Methodological problems associated with conservation testing, related to task validity, horizontal decalage, and inference criteria, are isolated and discussed. Conservation tasks generally are analyzed, and four elements essential to task validity are reduced. On the basis of the theoretical discussion, an unequivocal testing procedure for volume conservation employing nonverbal responses and objective criteria is designed. (Dr. Hobbs is Assistant Professor in the Department of Science Education, University of British Columbia.)

The ultimate concern in this article is the design of a valid, objective test of volume conservation. First, a number of methodological problems endemic in Piaget-related research will be isolated and discussed. Conservation tasks specifically will then be considered and finally, a reliable procedure for testing volume conservation, using valid objective criteria, will be proposed.

The central developmental predictions of Piaget's theory concern sequences and synchronisms. That is, certain abilities are supposed to emerge in sequences, while others are supposed to emerge concurrently. The theory posits that these sequences and synchronisms are not coincidental, do not imply innate structures, and do not result purely from learning experiences. Instead, they are viewed as the inevitable result of certain laws of development, with development seen to be the progressive construction of equilibrated systems of operations. The nature of the theory being so, related research usually involves one or more of four kinds of activity: (a) classifying subjects according to developmental stage or "presense" of a particular ability. (b) empirically verifying the sequences and synchronisms, (c) setting up scales of cognitive growth on the basis of the sequences, and (d) articulating the stage theory.

The first kind of activity, classifying subjects according to developmen-

tal level or performance on Piagetian tasks, may be done in order to characterize some defined population, or to compare populations in cross-cultural studies.

The second kind of activity, verifying invariant sequences or synchronisms, is necessary to establish the empirical base of Piagetian theory. According to Brainerd (1973), "it is the hypothesized *correspondences* and *sequences* which are the crucial empirical predictions [of Piagetian theory]." Studies involved in this kind of activity usually require the classifying of subjects in cross-sectional or longitudinal surveys. Some employ training experiments, which also involve classifying subjects according to performance on Piagetian tasks, after and often before training procedures.

The third kind of activity, setting up scales of cognitive growth, as exemplified by the work at Temple University (McLaughlin, Bentler, & Stephens, 1970), involves assuming that invariant developmental sequences exist. A large number of Piagetian tasks are standardised and placed on an ordinal scale. A subject's position on the scale may then be determined by the classification of his behavior on the tasks.

The fourth kind of activity, articulating the stage theory, is necessary because so far the sequence of development is worked out only in coarse detail. The fine structure remains to be established (e.g., Somerville, 1974). Another problem of articulation, one with which the Genevan group is occupied, is that of the mechanism by which subjects pass from one developmental stage to the next (Piaget, 1971). Both kinds of problem may require training experiments, which in turn involve classifying subjects.

Thus all four streams of theory-related research involve classifying subjects according to performance on Piagetian tasks. All Piagetian tasks may be reduced to the following simple paradigm: (a) *S* is exposed to a complex set of stimuli—"the task," (b) *S*'s behavior is observed, and (c) from this behavior, developmental stage, psychological processes, or thought structures are inferred in *S*. Three fundamental methodological problems arise in this paradigm. The first may be termed "task validity," the second is concerned with the phenomenon of horizontal decalage, the third is that of inference criteria. These will now be discussed seriatim.

Task Validity

The term "validity," applied to a Piagetian task, has two meanings. The first refers to the validity of the Piagetian interpretation of subjects' performance on the task. The second refers to the validity of applying that same interpretation to a variation of the task. The act of labelling a particular task as, for example, a volume conservation task, is actually an interpretation of subjects' behavior on the task. To give the same label to a variation of the task is to make a judgment about the equivalence of the task and variation. Such a judgment may not be valid because subjects' behavior on the variation may have a different interpretation in the same theory. Thus the designing of a task necessarily involves interpretation of the theory. If the experimenter is assuming the validity of the Piagetian interpretation of a given task, he must be careful not to apply that interpretation to a variation of the task without justification in terms of the theory. Classification of subjects in experiments which lie within the framework of the theory

will have little meaning if based on tasks which in effect reinterpret the theory. On the other hand, if the experimenter is challenging the validity of the Piagetian interpretation of a task, he must be careful not to apply the Piagetian interpretation to a variation of the task, unless that interpretation can be justified in terms of the theory. Otherwise he will be testing the validity of his own reinterpretation of the theory rather than the theory itself. This is not to exclude the legitimacy of redesigning tasks on the basis of conscious reinterpretation of the theory. Later in this article a number of elements essential to valid testing for conservation will be derived from an analysis of conservation tasks in terms of Piagetian theory.

Horizontal Decalage

The phenomenon of horizontal decalage complicates all four kinds of research activity, but especially those of verifying sequences and synchronisms, and setting up scales of cognitive growth. Simply put, the problem is that different tasks, which may be equally valid tests for the same developmental level, intellectual ability, or thought structure, may not be equally difficult. Determining whether or not two given abilities are acquired in fixed order or synchronously obviously requires that the tasks used to infer these abilities are equally difficult. Flavell and Wohlwill (1969) have proposed a competence-performance model to meet this problem. The practical use of this model must be delayed however until the values of parameters representing task difficulty can be established empirically. The model also has the disadvantage that it requires the inference of an "ideal" competence which can never be measured directly. A practical course of action is to use a number of tasks for the same ability, together with the strategy of classifying some subjects as transitional cases. This course of action will be justified later in the article for the case of conservation.

Inference Criteria

The problem of inference criteria is that of deciding which aspects of the subject's behavior on the task constitute necessary and sufficient criteria for inferring the developmental level, ability, or thought structure for which the task is a valid test. If the criteria are necessary but not sufficient, type I errors will be made; while if they are sufficient but not necessary, type II errors will be made. Brainerd (1973) appears to have resolved the issue of whether judgments alone should constitute criteria, or whether judgments and explanations should be required. He showed that if care is taken in the design of the tasks, a correct judgment alone can provide a necessary and sufficient criterion, whereas the requirement of a satisfactory explanation constitutes a sufficient but not a necessary criterion. Braine (1970) also advocates the use of nonverbal tests. The approach taken in this article in the case of conservation testing will be to emphasize the necessity of the use of objective operational measures of the quantities in question, whether verbal or nonverbal tests are used.

Conservation Tasks

Before proceeding to analyse conservation tasks generally, the need for concern with these methodological problems will be further emphasized by

illustrating the wide variation in volume conservation testing methodologies which have been used by different researchers. Four specific examples will be described: procedures used by Inhelder (1968), Elkind (1961), Goodnow and Bethon (1966), and Phillips (1971).

In the Inhelder procedure, subjects were shown three different transformations of a clay ball. In each case they were questioned about the room occupied by the bodies in water, and were asked to predict water levels after immersion of the transformed object in water. Subjects were cross-questioned by the experimenter in a face-to-face situation. Classification of the subject was made by the experimenter on the basis of all of the subject's responses. This famous "clinical technique" has been well described by Flavell (1963).

In the Elkind procedure, subjects were shown only one transformation of a clay ball. Subjects were tested in groups and recorded their own responses to a series of prepared questions with pencil and paper. Subjects were asked for five responses: they were asked (a) to say whether the two balls occupied the same amount of room or space, (b) to predict whether the two objects would occupy the same amount of room or space should one be rolled out into a sausage, (c) to judge after the transformation was actually carried out whether the two configurations did occupy the same amount of room or space, (d) to explain their responses, and (e) to predict whether each configuration would cause equal rises in water level upon immersion in identical containers of water. Subjects were classified as conservers only if they gave satisfactory responses in every case.

In the Goodnow-Bethon procedure, subjects were shown only one transformation of a clay ball. They were shown the water level rise which occurred when the untransformed ball was immersed in a glass of water and were required to predict what the level would be should the final configuration be immersed in an identical glass of water. Subjects were then asked to explain their responses. They were interviewed individually with a prepared set of questions. Subjects were classified as conservers or non-conservers according to their predictions of the water level, and conservers were further classified according to their explanations of equality of levels: subjects who referred to weight in their explanations were included in one category and not in the other.

In the Phillips procedure, subjects were shown pictures of the initial configurations, the final configurations and the transformations. The initial and final configurations were different arrays of one-inch metal cubes, immersed in water or in a medium of small solid cubes. Subjects were questioned individually and required to judge whether the immersing medium would rise to the same level in its container for the two different arrays of cubes, or for the same array in two different positions in the container. Subjects were required to explain their answers and were classified as conservers only if both correct judgments and correct explanations were given.

Other procedures, different from each of these, could be cited, but these four suffice to illustrate the diversity among procedures which purport to test for volume conservation. The theoretical analysis which follows will show that each of these procedures is subject to one or more sources of removable systematic error.

Volume conservation may be simply defined. It is the recognition that a given mass of material which undergoes a change of shape (without change in density) does not change in the volume it occupies. The difficulty lies in devising a test which will allow the experimenter to decide unequivocally whether such recognition may be inferred to a given subject. Volume conservation tests are usually variations of a procedure first devised and used by Inhelder (Piaget & Inhelder, 1968). Some experimenters (e.g., Lovell & Ogilvie, 1961) have used procedures similar to ones described in *The Child's Conception of Geometry* (Piaget, Inhelder, & Szeminska, 1960). Both kinds of procedure are forms of various quantity conservation tests used by Piaget over the years to test for conservation of substance, length, weight, and so on. Conservation testing generally will now be subjected to a theoretical analysis.

There are three essential steps in an orthodox quantity conservation test. These are:

1. Two identical objects are shown to the subject in some initial configuration—e.g., two balls of clay.
2. One of the objects is transformed into a new configuration as the subject observes—e.g., one of the balls is rolled out into a sausage.
3. The subject is required to compare the transformed object with the untransformed object in terms of some operational measure of the quantity being tested—e.g., the subject is asked if the ball and sausage will weigh the same on a balance, or will raise the levels of water in two identical containers to the same height.

These three steps in a conservation test may be characterized further by four elements which are essential to task validity. The first element is *presence of both the final configuration and the initial configuration for comparison by the subject*. Elkind (1967) has shown the importance of starting in the first step with two objects, one of which remains unchanged throughout. Should only one object be shown, then a subject may appear to conserve merely by recognizing that it is the same object after the transformation as it was before. In other words the subject recognizes the identity of the two configurations. Mere identity, otherwise called "pseudo conservation" (Piaget, 1967) can occur when one essential aspect of the conservation problem is missing: immediate perceptual difference between the two configurations. When there is no initial configuration for comparison the subject has to imagine what the initial configuration was, and as Elkind (1967) puts it, through "memory falsification" the subject may arrive at pseudo conservation. Similarly, the final configuration must be present in the final comparison. Piaget (Piaget, 1967; Piaget & Inhelder, 1971) has demonstrated the existence in subjects of "false anticipatory images." Thus if a subject who recognizes identity of body in a transformation, his memory of the initial configuration or his anticipatory image of the final configuration may be such that the observed and imagined configurations seem identical.

The second element is *demonstration of the transformation*. Without the transformation, the test is not of conservation but of estimation of the physical quantity. According to Elkind (1967), estimation is a more difficult task than conservation. Moreover logical necessity, which is characteristic of conservation, is not possible for such an estimation. Phillips (1971) has

suggested that actual demonstration of the transformation may not be essential. He tested two groups of subjects for conservation, the transformations being demonstrated to subjects in one group, and described with the aid of pictures to subjects in the other group. Phillips concluded no significant difference ($p > .05$) between the proportions of conservers in the two groups. Nonetheless a close inspection of his results shows that higher proportions of conservers were found when actual objects were shown than when pictures were shown. In fact for all of the conservation tests combined (quantity, interior and exterior volume) the difference may be taken to be significant ($F = 2.60$, $df = 1,118$, $p < .20$). Thus to use pictorial presentation of the objects and transformations may result in type II errors.

The third essential element of a conservation test is *use of a variety of transformations*. Subjects who are in a stage of transition from nonconservation to conservation find some transformations more difficult than others. This problem comes under the heading of "horizontal decalage." Piaget (1971) has compared it with "friction," while the analogue used by Flavell and Wohlwill (1969) is "noise." These analogies imply the existence of an ideal or pure "motion" or "signal" which could be found in principle by creating measuring situations where the "friction" or "noise" is reduced to zero. Thus there is postulated an ideal or pure ability. Piaget's approach is to consider the problem to be in principle quite intractable, but not at all crucial—something of a nuisance. Flavell and Wohlwill quite correctly recognize the extreme difficulty the problem poses for the work of verifying the "sequential patterns of cognitive structure." They have proposed a competence-performance model, which involves an equation for predicting the probability of success on a given task by a given subject. The equation contains parameters representing both the "noise" and the "pure signal" (that is, the ideal or "true" competence of the subject). A fundamental objection to this model is that it involves a pure or ideal ability which can never be directly measured; but even if this is acceptable, the model cannot be put to practical use until the chore of generating empirical data to define its parameters is in hand. In the meantime, the precaution should be taken to use a range of transformations in conservation testing. Subjects who conserve on some transformations but not on others may then be classed as transitional cases. If a dichotomy of subjects is required, the transitional cases may be classed as conservers whenever type II errors are considered to be more costly, and as nonconservers whenever type I errors are considered to be more costly.

The fourth element is the *use of an objective operational measure for the quantity to be conserved*. The initial and final configurations are compared in terms of this objective, operational measure. Some examples of operational measures are: for length—distance to be walked, area—grass for cows to eat in a field, weight—balance with another object on an equal-arm balance, substance—amount to eat or drink.

The experimenter cannot be sure that the subject understands the question if it is asked with the use of terms as "same volume" or "just as heavy," because in their early stages of development, quantity concepts are gross or undifferentiated (*global*). Verizzo (1970) found that terms such as "mass," "weight," and "volume" were virtually synonymous even for very bright adolescents. Furthermore, the experimenter should be careful not to

induce the subject to think only in perceptual terms of the quantity being tested. In Piagetian theory, the achievement of conservation of a physical quantity means that the subject has progressed from (a) notions of the quantity in terms of the egocentric or perception-based effects by means of which he came to be aware of it, to (b) understanding of the quantity in terms of objective interaction between bodies (Piaget & Inhelder, 1968, p. 59). Some experimenters ask comparison questions such as: "Now is the sausage heavier than the ball, or lighter, or the same?" without reference to a weighing balance. To do this is to lead the subject to rely on egocentric notions of the weight (force on the hand or "hardness"). For instance when Hyde (1970) tested weight conservation with the ball-sausage transformation, she actually asked subjects to heft the ball and sausage in their hands, thus inducing them to answer in terms of their perceptual judgments. Some of Hyde's nonconservers of weight might have conserved if the question had been posed using a balance scale. In addition, according to Griffiths, Shantz, & Sigel (1967) there may be problems associated with the use of relational terms, such as "more," "less," and especially "same," at least for preschool children. Some experimenters have resorted to "nonverbal" methods (e.g., Braine, 1968; Mermelstein & Shulman, 1967). In fact the experimenter who devises nonverbal methods is forced to use objective operational measures for the quantities in the conservation tests. Apparently, the use of nonverbal methods has indicated that reliance on the verbal mode in conservation testing may increase the probability of type II error. The preferred course of action is to use nonverbal methods whenever these can be devised and administered readily, but the nonverbal tests usually require either training sessions or complicated apparatus or both. Whenever verbal methods are more feasible, the experimenter must remember that type II errors may be made, and must ensure that the quantity to be conserved is referred to in terms of an objective operational measure, avoiding such words as "longer," or "smaller," or "weighs the same."

To recapitulate, there are four elements in the three-step conservation test which are essential to task validity:

1. presence of both the initial and final configurations,
2. demonstration or at least description of the transformation,
3. use of variety of transformations, and
4. use in the final comparison of an objective, operational measure of the physical quantity to be conserved.

The first and second of these elements place constraints upon the testing procedure. The third and fourth allow for some variation in procedure in terms of the transformations used and the specific operational measure of the physical quantity employed. The possible range of variations is limited by the requirement to avoid reinterpreting the theory. Thus specific design of a conservation task involves selection of the transformation to be used and selection of an objective operational measure of the quantity to be conserved. This will now be done for the case of volume conservation, starting with transformations.

Transformations

Piaget and his collaborators have used three kinds of transformation in volume conservation tests: rearrangement of arrays of solid cubes, deforma-

tion of spheres of clay (plasticene), and the dissolving of sugar in water. In the first method (Piaget et al., 1960), the subject has to recognize that the volume or room occupied by a stack of cubes (say $4 \times 6 \times 10$) does not change when the cubes are restacked in a different way (say $6 \times 5 \times 8$). In the second method (Piaget & Inhelder, 1968; Inhelder, 1968), a sphere of clay is either (a) rolled out into a "sausage," (b) flattened out into a "pancake," or (c) broken into several pieces. In the third method (Piaget & Inhelder, 1968; Inhelder, 1968), subjects are shown sugar lumps being placed in water and subsequently dissolving. Subjects note the rise in water level on addition of the lumps, and are then required to say whether the increased level will be maintained as the sugar dissolves.

According to Piaget, "There is a good deal to be gained in avoiding the use of discontinuous elements . . . with older children who would otherwise be able to solve the problem by arithmetical correspondence without considering the volume as such" (Piaget et al., 1960, p. 356). This means that the restacking of cubes method may be prone to type I errors, especially with older children. On the other hand there is the chance of type II errors if a very difficult or unusual transformation is used, because to some extent a volume conservation test involves understanding the transformation. To understand the transformation requires the construction of a grouping of physical or infralogical operations to represent the transformation (Piaget & Inhelder, 1968). Such groupings involve "atomistic schemes," in which shape changes are represented by movements of elementary pieces of the object from one place to another in the mass (these movements being the physical operations). There are physical transformations in which volume is not conserved (e.g., freezing, or mixing of some liquids), and obviously mature volume conservers are readily able to assimilate such transformations. Therefore two achievements may underlie conservation of volume in a given transformation: (a) volume conservation as defined above, which involves differentiation of volume of a body from other properties (shape, surface area, etc.) and (b) understanding of the transformation, which involves construction of a grouping of physical operations. Type II errors may occur if the subject has the first ability (volume conservation per se) and not the second (understanding of the transformation). Consequently very difficult transformations should be avoided. To understand the deformation of clay requires no more than an atomistic scheme for envisaging the transfer of parts of the clay from one location in the mass to another. This kind of scheme is very similar to the scheme of additive compensations, which is acquired in the concrete-operational period. Volume conservation however requires the scheme of multiplicative compensations, which is acquired in the formal-operational period (Inhelder & Piaget, 1958). This means that for deformation of clay, the kind of type II error just described is low since the understanding of the transformation requires an atomistic scheme which is accessible to thinking at an earlier developmental level than at least one scheme (multiplicative compensations) necessary for volume conservation. On the other hand, the atomistic scheme necessary to understand dissolution of sugar involves the separation from one another into the water of invisible particles of sugar, *each with a fixed volume* and such that the total volume of the sugar lump is equal to the sum of the volumes of these individual particles. Therefore,

this scheme itself presupposes volume conservation, so that the kind of type II error just described may be more likely for this transformation than for deformation of clay. Furthermore, there is in fact a contraction of volume as the sugar dissolves (even though very small) which might make the test suspect for older children who might be aware of this (from science classes perhaps).

Objective Operational Measures

Careful investigation into the relative difficulties of various transformations may provide better understanding of the development of volume conservation. Uzigiris (1964) found that difficulty varied as a function of stimulus material, while an experiment by the writer (1973) indicated that the relative magnitude of the change in surface area may be the important factor determining difficulty. In the absence of any firm evidence about the relative difficulty of transformations, and in view of the preceding analysis, a prudent course would be to employ the three deformation of clay transformations or minor variations of these.

Consider now the operational measure to be used in the comparison of volumes of the initial and final configurations. Most experimenters who have employed an operational measure for volume have, like Piaget and Inhelder, used the rise in level which occurs when an object is placed in a container of water. There is a problem with this measure in that the subject must not only conserve volume, but also realize that the volume and not some property of the body determines the amount of change in the water level. Piaget's research has shown that in the concrete-operational period children believe the rise in the level to be caused by the object's weight (Piaget & Inhelder, 1968). Therefore if volume conservation is tested by asking only if the initial and final configurations would cause the same rise in water level, a subject may predict equality of the rises in level on the basis of conservation of weight rather than of volume. An experiment by the writer (1973) showed that up to 12% of junior high school children may appear to conserve volume by conserving weight when the test involves questions about water levels only, without references to volume or room occupied. This represents about 16% of those who would be classed as conservers on such a test.

In Vinh-Bang and Inhelder's experiments, from 10% to 30% of subjects (depending on the age group) explained equality of water levels in the volume conservation test by action of weight (Piaget & Inhelder, 1968, p. xiii; Inhelder, 1968, p. 317). The reader must decide whether these subjects should be classed as conservers. Goodnow and Bethon (1966) gave two sets of results for volume conservation acquisition; one including all subjects who anticipated equality of water level, the other excluding subjects who explained this equality by reference to equal weight. Of a group of "average ability" eleven-year-old boys, 56% were in the first category, 31% in the second.

To be so ambiguous about the results of a volume conservation test is both undesirable and unnecessary. Subjects who correctly predict equality of water levels merely by conserving weight should not be classed as conservers of volume. To avoid the possibility of type II errors which might then occur, efforts should be made to ensure that the subject understands

the displacement of water level to be a measure of space or room occupied by the body immersed. In other words, steps should be taken to make sure subjects are aware that the question being asked refers to the volume or room occupied by the two bodies. When this precaution is taken, subjects who persist in explaining equality of water levels in terms of weight should not be classed as volume conservers.

In fact care was taken in the original volume conservation experiments to direct subjects' attention to the volume occupied by the bodies in the test. Piaget and Inhelder describe the procedure as follows:

We decided after trying various methods to have volume estimated by means of the space occupied by the ball and by its successive forms in a glass of water, this space being identified by the level of water. We mark the initial level of water with ink or with an elastic band. Then we ask, "If I put this ball in the water, will it take up room? Will the water rise or stay at the same level?" This first question is suggestive by design. Indeed we have established elsewhere that a certain proportion of subjects between five and eight years are ignorant of the phenomenon and are surprised to see the water rise upon immersion of the solid. On the other hand, we have been able to establish that most subjects before eight or nine years explain the rise in level by referring not so much to the volume as to the weight, which produces according to the child an upwards current. Now these two reactions are very interesting in the study of conservation of volume. . . .

But the matter at hand is to set the problem from the start in the field of volume. That is why we say to the child "take up room." [*prendre de la place*] After having established in the subject that the first ball made the level rise and having marked the second level, we present a second ball, similar to the first, asking how far it will bring up the level. Every child referred to in this chapter indicated the same second level and thereby showed himself capable of understanding the basic ideas in the problem. It is then and only then that the second ball is transformed into a sausage, disc, etc., or broken into pieces, with the question each time: "and now does it take up the same room in the water as the other ball? Where will the water rise to?" etc. (Piaget & Inhelder, 1968; writer's translation).

A further control is described by Piaget (Piaget & Inhelder, 1968, Introduction, pp. xvii-xviii): before comparison of the sausage and ball by the immersion in water, a metal ball is substituted for the original untransformed clay ball. Subjects who conserve weight but not volume would give an incorrect response in this test.

Now as careful as these precautions are, justification can be advanced for going beyond them. The last control mentioned is necessary to avoid classifying subjects as volume conservers when in fact they have done no more than conserve weight in the test. This control may however act as a "trap" for volume conservers who persist in explaining rise in water level in terms of the weight of the immersed object. To avoid this, the subject should be shown by explicit demonstration that (a) objects of equal volume but different weight cause equal rises in water level, and that (b) objects of equal weight but different volume cause different rises in water level. An objection to this procedure might be that volume conservation may be induced in subjects who would have been classified as nonconservers before the demonstration. Indeed, Vinh-Bang and Inhelder have shown that learning of volume conservation can occur rapidly in some subjects when they are shown immersion of ball sausage etc. in water. (Percentage of conservers went from 37% to 78% of eleven-year-olds after three

demonstrations—Inhelder, 1968, p. 318; Piaget & Inhelder, 1968, Introduction p. xviii.) An experiment by the writer (1973) indicated that discovery of the fact that the volume of the immersed body determines the rise in water level may precipitate acquisition of volume conservation in older subjects.

Inference Criteria

A reply to this objection necessitates consideration of the problem of inference criteria. Success or failure on a volume conservation test infers presence or absence of certain schemas, (e.g., proportionality and multiplicative compensations) and hence, at least the beginning of formal operations. To confine attention to what a subject actually *does* in a given test rather than to what he may be *able* to do under appropriate circumstances would be to allow temporary aberrations to be confused with absence of the theoretical underlying abilities. Consider a given subject who is unaware that an object's volume determines displacement of water, even though he has begun to construct formal operations. He conserves weight and is in possession of schemes of multiplicative compensations and proportionality, but yet may fail in a given test to conserve volume. Such failure would be merely a temporary mistake; such a subject would be found to acquire volume conservation very quickly in the face of a few simple demonstrations. To classify such a subject as a nonconservers would be to commit a type II error.

In the Piagetian clinical technique, the subject is given every chance to arrive at the correct answer. A temporary mistake (or for that matter a chance correct response) is not accepted at face value, but the subject is questioned further (Inhelder, 1968, pp. 344-346). Of course the clinical technique has its problems, associated with the subjectivity of classification criteria. A more objective, standardized procedure is sought here; one which can be administered to individual subjects or, if necessary, to groups of subjects recording their own responses with pencil and paper. In such a test, two kinds of response may be recorded: a subject's *judgment* of the equality of the two volumes, and his *explanation* of the equality. Brainerd (1973) argues that the thought structures to be inferred from conservation tests are, according to Piaget, relatively independent of language. Therefore adequate explanations are sufficient but not necessary conditions for the inference of a given thought structure. In addition (he argues) to use explanations as necessary criteria is to confine the response mode to the verbal, and this "unduly restricts the behavioral domain to which the theoretical construct (structure) applies." Thus the use of explanations as the sole criterion is liable on two counts to result in failure to infer the existence of a thought structure that does exist, i.e., explanations criteria "are subject to at least two sources of systematic type II error." Judgments, on the other hand, "from the standpoint of Piagetian theory . . . are not subject to any known sources of systematic error."

A Prototype Volume Conservation Test

Taking into account the requirements derived above, a test for volume conservation was designed and administered to a grade five class in an Edmonton (Canada) school. The test has three special features. First, a preparation of subjects was undertaken so as to ensure that they did

associate the objective operational measure—displacement of water level upon immersion of an object—with occupied volume. Second, the test avoided the use of relational terms such as “more,” “less,” or “same.” Third, a nonverbal response was required. An added but nonessential feature was the administration of the test by means of a videotape recording. Use of the video medium allowed the test to be administered to a whole class at one time, and avoided unwanted or uncontrolled subject-experimenter interactions. The videotape presentation was supplemented by actual “live” demonstration of the transformations. This avoided problems due to the distortion which accompanies the two-dimensional electronically-mediated TV picture.

The test will now be described as it would be administered to an individual subject by a “live” experimenter. Parenthetical notes will be used to describe those adaptations which were made for the videotape presentation.

1. *E* presents three identical glasses to *S*. The glasses are side by side and filled to the same level with water. The levels are marked by a black line on each glass. *E* points out that the levels are the same in each glass. The first glass remains unchanged throughout the experiment.

2. *E* displays five or six identical balls of plasticene and notes that they are all the same. He asks “What do you think will happen if I put this ball into this [second] glass? Will it take up any room in the water? Where will the water come to?”

3. When *S* has responded, *E* says “let’s try it,” and places the ball into the second glass. *E* says “See? The water came up to here.”

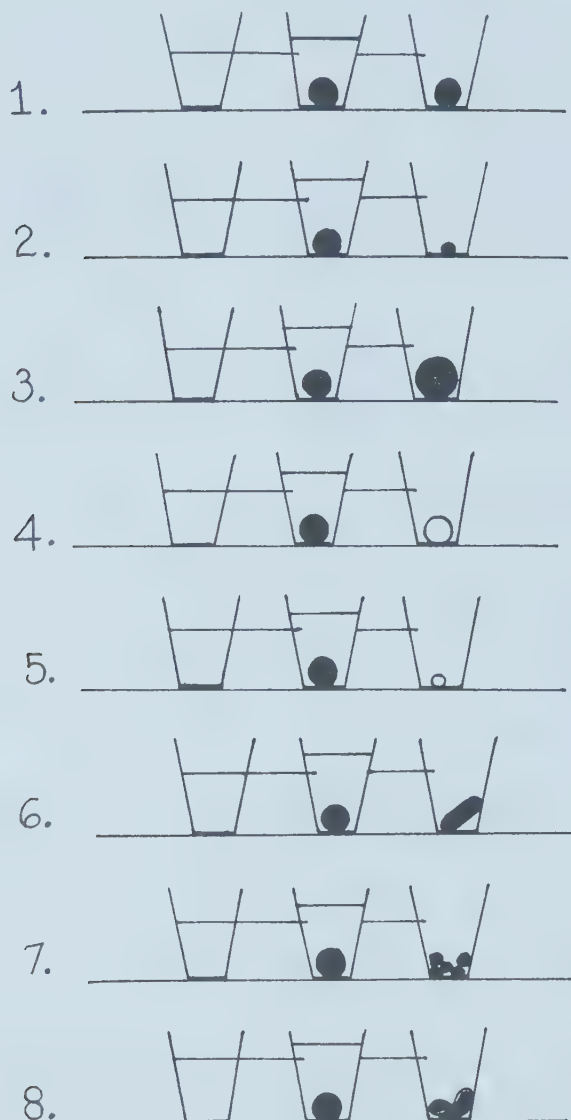
4. *E* takes another ball identical to the first and says “Now what if I put this ball into this [third] glass? How much room will it take up in the water? Where will the water come to? Show me.” *S* is required to respond by pointing to where he thinks the level will rise to in the third glass upon immersion of the ball. (In the videotape presentation, subjects were required to respond by marking in a level on a diagrammatic representation of the three glasses side by side, with the levels shown only for the first two. The answer sheet on which subjects responded is shown in Figure 1.) Note that the response is nonverbal in both cases, and that relational terms are not used in posing the question.

5. After *S* responds, *E* says “Let’s try it,” places the ball in the third glass, and says “See? The water came to here.”

6. Steps 4 and 5 are repeated using respectively (a) a larger plasticene ball, (b) a steel ball of the same size, and (c) a smaller lead ball. Each time the same question is asked: How much room will *this* ball take up in the water. If I put this ball into the glass, where will the water come to?” Each time, after *S* responds, the ball is placed into the third glass to show how much the water level rises. (In the videotape presentation, a fresh prefilled glass was used each time as the third glass.)

The procedure to this point represents the preparation of the subject. By responding incorrectly to the last question in the preparation, subjects indicate that they are persisting in associating the displacement of the water level with the weight of the immersed object. Such subjects should be classed as nonconservers to avoid type I errors, for these subjects may

FIGURE 1
RESPONSE SHEET FOR VOLUME CONSERVATION TEST



appear to conserve volume by conserving weight. On the other hand if type II errors are more costly, the preparation procedure may be repeated before classifying such subjects as nonconservers. With the preparation complete, *E* proceeds to the actual volume conservation tasks.

7. *E* presents two balls identical with the one immersed in the second glass. One is deformed into a cylinder or "sausage." *E* says "How much room will the sausage take up in the water? If I put the sausage in the [third] glass, where will the water come to?" *S* responds, but this time *E* does not actually immerse the transformed object. (In the videotape presentation, the transformation and comparison with the untransformed ball were demonstrated "live," interrupting the playback of the tape.)

8. Step 7 is repeated for two other transformations: (a) breaking the ball into four or five pieces, and (b) flattening the ball into a disc or "pancake."

Subjects who are successful in the preparation and respond correctly to all three volume conservation tasks are classified as conservers. Subjects

who are successful in the preparation and respond correctly to only one or two of the volume conservation tasks are classified as transitional cases. Subjects who are unsuccessful in the preparation or who respond incorrectly to all of the volume conservation tasks are classified as nonconservers. Success in the preparation is measured by response to the last question. (An objective method of judging whether the lines drawn by *S* were correct was used in the test run: the line drawn by *S* had to come within 0.5 mm of the correct level at some point to be judged correct. This was determined by centering a 1 mm thick line drawn on a transparent sheet over the correct level.)

In the grade five class tested ($N = 28$), 14 subjects were classified as conservers, 10 as nonconservers, and 4 as transitional cases. Only one subject responded incorrectly to all of the preparation questions and was thereby classified among the nonconservers. The average age of the subjects was 10 years 10 months. No problems arose in the administration of the test or in its scoring, and subjects indicated to the experimenter that they understood what they were required to do. The entire testing procedure took just over 10 minutes.

Summary

This volume conservation task provides an illustration of how methodological difficulties associated with Piagetian tasks can be overcome by means of careful design of the testing procedure based on analysis of the tasks. The methodological difficulties defined were concerned with task validity, horizontal decalage, and inference criteria. Consideration of validity led to the conclusion that design of the task has to be justifiable in terms of Piagetian theory, unless conscious reinterpretation of the task is intended. In conservation tasks, this means that the transformation should be demonstrated, and that the initial and final configuration have to be simultaneously seen and compared by the subject. Consideration of horizontal decalage led to the conclusion that a variety of forms of the task should be employed. In conservation tasks this means that a number of transformations should be demonstrated, and that very difficult transformations should be avoided. Consideration of inference criteria led to the conclusions that nonverbal methods should be used wherever feasible, that judgments alone represent sufficient criteria, and that in conservation tasks, objective operational measures should always be used for the quantity to be conserved.

Assistance in preparation of this article was provided by Dr. C. Brainerd and Dr. L. Stewin of the University of Alberta.

References

- Beilin, H. The training and acquisition of logical operations. In M. F. Roszkopf, L. P. Steffe, & S. Taback (Eds.), *Piagetian cognitive-developmental research and mathematical education*. Washington: National Council of Teachers of Mathematics, 1971.
- Braine, M.D.S. The ontogeny of certain logical operations: Piaget's formulation examined by nonverbal methods. In I.E. Sigel & F.H. Hooper (Eds.), *Logical thinking in children*. New York: Holt, Rinehart & Winston, 1968.

- Braine, M. D. S. Piaget on reasoning: A methodological critique and alternative proposals. In Society for Research in Child Development. *Cognitive development in children—Five monographs of the Society for Research in Child Development*. Chicago: The University of Chicago Press, 1970.
- Brainerd, C. J. Continuity and discontinuity hypotheses in studies of conservation. *Developmental Psychology*, 1970, 3, 225-228.
- Brainerd, C. J. Judgements and explanations as criteria for the presence of cognitive structures. *Psychological Bulletin*, 1973, 79, 172-179.
- Elkind, D. Quantity concepts in junior and senior high school students. *Child Development*, 1961, 32, 551-560.
- Elkind, D. Piaget's conservation problems. *Child Development*, 1967, 38, 15-27.
- Flavell, J. H. *The developmental psychology of Jean Piaget*. New Jersey: D. Van Nostrand, 1963.
- Flavell, J. H., & Wohlwill, J. Formal and functional aspects of cognitive development. In D. Elkin & J. H. Flavell (Eds.), *Studies in cognitive development: Essays in honor of Jean Piaget*. New York: Oxford University Press, 1969.
- Goodnow, J. J., & Bethon, G. Piaget's tasks: The effects of schooling and intelligence. *Child Development*, 1966, 37, 537-582.
- Graves, A. J. Attainment of conservation of mass, weight, and volume in minimally educated adults. *Developmental Psychology*, 1972, 7, 223.
- Griffiths, J. A., Shantz, C. A., & Sigel, I. E. A methodological problem in conservation studies: The use of relational terms. *Child Development*, 1967, 38, 841-848.
- Higgins-Trenk, A., & Gaite, A. J. H. Elusiveness of formal operational thought in adolescents. *Proceedings of the 79th Annual Convention of the American Psychological Association*, 1971, 6, 201-202.
- Hobbs, E. D. Adolescents' concepts of physical quantity. *Developmental Psychology*, 1973, 9, 431.
- Hyde, D. M. G. *Piaget and conceptual development*. London: Holt, Rinehart & Winston, 1970.
- Inhelder, B. *The diagnosis of reasoning in the mentally retarded*. New York: John Day, 1968. Trans. from the French *Le diagnostic du raisonnement chez les débiles mentaux*. (2nd ed.) Neuchatel: Delachaux et Niestle, 1963. (1st ed. 1943).
- Inhelder, B. N., & Piaget, J. *The growth of logical thinking from childhood to adolescence*. New York: Basic Books, 1958. Trans. from the French *De la logique de l'enfant à la logique de l'adolescent*. Paris: Presses Universitaires de France, 1955.
- Keller, H. R., & Hunter, M. L. Task differences on conservation and transitivity problems. *Journal of Experimental Child Psychology*, 1973, 15, 287-301.
- Lefrancois, G. R. A treatment hierarchy for the acceleration of conservation of substance. *Canadian Journal of Psychology*, 1968, 22, 277-284.
- Lovell, K., & Ogilvie, E. The growth of the concept of volume in junior school children. *Journal of Child Psychology and Psychiatry*, 1961, 1, 191-202.
- McLaughlin, J. A., Bentler, P. M., & Stephens, B. Monotonicity analysis of Piagetian assessments of reasoning. *Proceedings of the 78th Annual Convention of the American Psychological Association*, 1970, 5, 283-284.
- Mermelstein, E., & Shulman, L. S. Lack of formal schooling and the acquisition of conservation. *Child Development*, 1967, 38, 39-52.
- Phillips, D. J. The development of the concept of displacement volume: A hierarchical model and its partial testing under two methods of presentation. *Journal of Research in Science Teaching*, 1971, 8, 9-19.
- Piaget, J. Untitled article. *Contemporary Psychology*, 1967, 12, 532-533.
- Piaget, J. The theory of stages in cognitive development. In D. R. Green, M. P. Ford, & G. B. Flamer, (Eds.) *Measurement and Piaget*. New York: McGraw-Hill, 1971.
- Piaget, J., & Inhelder, B. *Le développement des quantités physiques chez l'enfant*. (3rd ed.) Neuchatel: Delachaux et Niestle, 1968. (1st ed. 1941).

- Piaget, J., & Inhelder, B. *Mental imagery in the child*. London: Routledge & Kegan, Paul, 1971. Trans. from the French *L'image mentale chez l'enfant*. Paris: P.U.F., 1966.
- Piaget, J., Inhelder, B., & Szeminska, A. The child's conceptions of geometry. London: Routledge & Kegan, Paul, 1960. Trans. from the French *La géométrie spontanée de l'enfant*, 1947.
- Sigel, I. E., & Hooper, F. H. *Logical thinking in children*. New York: Holt, Rinehart & Winston, 1968.
- Somerville, S. C. The pendulum problem: Patterns of performance defining developmental stages. *British Journal of Educational Psychology*, 1974, 44, 266-281.
- Uzigiris, I. C. Situational generality of conservation. *Child Development*, 1964, 35, 831-841.
- Verizzo, O. Conceptions of conservation and reversibility in children of very superior intelligence. *School Science and Mathematics*, 1970, 70, 31-36.

S R. R U T H E A G A N

The University of New Brunswick

An Investigation Into the Relationship of the Pausing Phenomena In Oral Reading and Reading Comprehension

This study attempted to investigate whether misuse of the pause in oral reading was related to the silent and oral reading comprehension abilities of young children.

Although the pausing phenomena may be considered as one product of the reading act, in this study it was considered to be an overt manifestation of how children might be processing visual input while reading silently and orally. The results of the study revealed that the pausing phenomena in oral reading were consistently and significantly able to discriminate between silent reading ability groups.

However, while the "processes" of silent and oral reading seemed to be similar, the "products" (or comprehension) of silent and oral reading did not appear to be related. Further, those children who were the better silent readers seemed to be impeded by the necessity of reading orally. Nevertheless, oral reading appeared to assist comprehension for the younger and less proficient readers. (Sr. Eagan is Associate Professor in the Faculty of Education. The University of New Brunswick.)

This study was an attempt to determine the nature of the pausing phenomena in oral reading, and whether these phenomena are possible overt features of the verbal processing entailed in reading orally, and also in reading silently, by relating them to the silent and oral reading comprehension abilities of the subjects.

In this study, the visual input, or orthographic string, was not organized for the child. It was assumed that how the child used the pause in his oral reading indicated how he was organizing the visual input. By investigating the number, length and placement of pauses as the child used these devices in oral reading, it was hoped that cues would be obtained to indicate whether the child was organizing the reading material into meaningful word groupings, or whether by pausing within such groups, he was disrupting the linguistic and cognitive unit.

The unique feature of this study was that the method by which the pausing phenomena were measured was a completely objective and repeatable procedure.

Purpose of the Study

The purpose of the study was to investigate the relationship between pauses made by children reading orally, and their reading comprehension ability, both silent and oral, and in particular, to determine whether the use, or rather misuse of the pause (by disrupting a linguistic unit) was related to silent reading comprehension groups, and oral reading comprehension scores.

The study also attempted to test the assumption that young children learning to read use oral reading processes that are similar to those they use when reading silently.

Definition of Terms

The following are some of the key definitions used in this study:

Pause: in this study, the term "pause" referred to the interruption or termination of the voice stream in oral reading. It was measured by an *Esterline Angus Speed Servo* machine, in milliseconds (See Appendix A).

Criterion Pause: since pause duration shows marked variability across oral readers, a "criterion pause" was established for each subject in the study. This was accomplished by calculating the mean length of the between-word pauses produced by subjects in their oral reading of a set of syntactic constituents (defined below).

Significant Pause: a significant pause exceeded the duration of an individual's "criterion pause" by a critical amount. This critical interval was precisely determined by means of pilot work. For purpose of this study, a significant pause was 2.5 times longer than the criterion pause.

Syntactic Constituent: this referred to the lowest major constituent as defined by Latham (1972). It may be formed from all labelled nodes in a surface structure tree. These constituents are found by locating those nodes which are immediately above the lexical nodes and deciding whether or not there are sister-nodes to the lexical nodes. If there are no sister-nodes to a specific lexical node, then the node immediately dominating that lexical node is a lowest major constituent. If there are sister-nodes to a specific lexical node, then the lowest major constituent associated with that node is the node which immediately dominates all sisters of the lexical node in question (See Appendix B).

Word Recognition: referred to the ability of the child to look at a word and pronounce it aloud correctly, as determined by the judgment of the investigator.

Theoretical Framework

Pause can serve two obvious purposes in the processing of language: (1) it serves to separate the string into meaningful units for analyzing content (Osgood, 1954; Garrett, Fodor, & Bever, 1966; Neisser, 1967;

VanUden, 1970; Bond, 1971) and (2) it facilitates the storage of items in short-term memory (Dale, 1964; Miller & Isard, 1963; Wickelgren, 1965, 1969; Neisser, 1969; O'Connell, Turner and Onoska, 1968; Smith, 1971). Any disturbance in these two uses of the pause in the processing of written language by children learning to read, may seriously impair their comprehension.

Research with adults has indicated that the grouping of words in reading orally is very consistent (North & Jenkins, 1951; Kolars, 1970; Cromer, 1970; Wilkes & Kennedy, 1970; Martin, Kolodziej, & Genay, 1971; Goldman-Eisler, 1968). Pause in oral reading by adult fluent readers always occurred at grammatical boundaries. Since a group of words, separated by pauses, has a certain cohesiveness, it follows that if this group is meaningful, it can more readily be stored as a unit in the short-term memory. It also follows that if this group lacks cohesiveness because of pausing within the group itself, then comprehension is likely to deteriorate.

Since the major goal of reading is comprehension, it is reasonable to suppose that the perceptual and cognitive processing of the material would be facilitated if the units stored for processing corresponded to the major sense groupings in the material being read. It seems then, that if the words of a syntactic constituent group were stored in the short-term memory as one unit, that comprehension would be facilitated. If, on the other hand, pausing occurs within the syntactic constituent group, it could be assumed that comprehension would be adversely affected.

Reading begins with the graphic input and ends with meaning as output. What happens between input and output has been hypothesized by many researchers, and is often called the reading "process". Since one cannot get inside a child's head to find out what is happening there, one way of attempting to investigate the reading process is to examine the output, or the "product". This product can be analyzed in two ways: (1) by examining the child's comprehension of what he has read, and (2) by analyzing the oral output in an oral reading performance. A child learning to read is, for the first time in his linguistic experience, exposed to the syntactic structure of the language from which the prosodic structure (one feature of which is pause) is absent. The present study is concerned with the use of the pause as a behavioral manifestation of how the young reader may be using the surface structure relationships to comprehend the message.

The models of beginning silent reading (Figure 1) and beginning oral reading (Figure 2) both indicate that the child recodes to aural input and/or oral output before he decodes the meaning, or comprehension. If during this intermittent step between graphic input and meaning the young beginning reader loses the cohesiveness of the word group because he pauses within this group, then such a disjunction may interfere with comprehension.

FIGURE 1
AN ADAPTATION OF GOODMAN'S 1970 MODEL
TO ILLUSTRATE EARLY SILENT READING

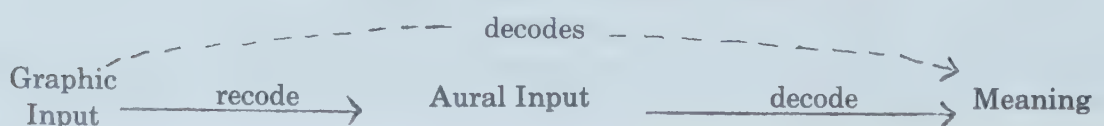
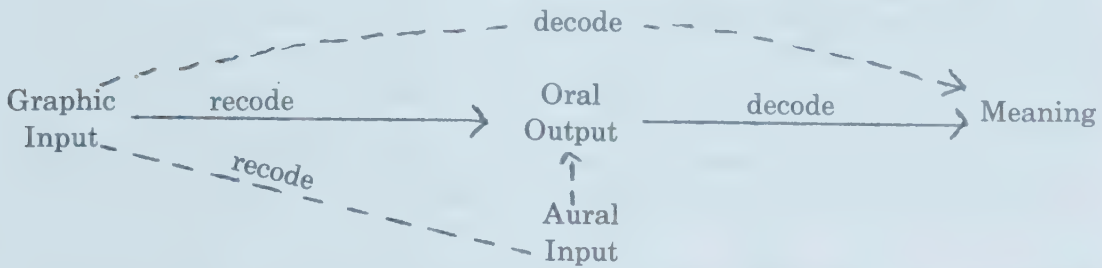


FIGURE 2
INVESTIGATOR'S MODEL OF EARLY ORAL READING



If, as assumed in this study, the silent and oral reading processes of young beginning readers are comparable (see Figures 1 and 2), then an analysis of how these young children use timing cues (pauses) in their oral reading, and the relationship of the use of these cues to comprehension (whether silent or oral comprehension), may provide some insights into some of the learning to read processes.

Hypotheses

Following are the main null hypotheses tested by this investigation:

1. That average readers, above-average readers and below-average readers do not differ significantly in the percentage of total reading time spent in pausing while reading orally,
 - a. when the effect of inadequate word recognition ability is eliminated from the data;
 - b. when the effect of inadequate word recognition ability is included in the data.
2. That average readers, above-average readers and below-average readers do not differ significantly in the period of time spent pausing within syntactic constituents while reading orally,
 - a. when the effect of inadequate word recognition ability is eliminated from the data;
 - b. when the effect of inadequate word recognition ability is included in the data.
3. That average readers, above-average readers and below-average readers do not differ significantly in the number of pauses made within syntactic constituents while reading orally,
 - a. when the effect of inadequate word recognition ability is eliminated from the data;
 - b. when the effect of inadequate word recognition ability is included in the data.
4. That there is no significant relationship between (1) the percentage of pause time used by children reading orally, (2) the period of time spent pausing within syntactic constituents while reading orally, (3) the number of pauses made within syntactic constituents by children reading orally, and
 - a. their oral reading comprehension scores.
5. That there is no significant relationship between (1) the percentage of total reading time spent in pausing, (2) the period of time spent in pausing within syntactic constituents, (3) the number of pauses made within syntactic constituents, and
 - a. Auditory Memory Span for Digits Forward

- b. Auditory Memory Span for Digits Backward
- c. Visual Memory Span for Letters, and
- d. Intelligence Quotients.

The Experimental Design

Sample

A stratified random sampling procedure was utilized to select the sample. Twelve average, twelve above-average, and twelve below-average readers were selected from each of grades two and three, from four schools, making a total of 72 children—six cells consisting of 12 children each. In each cell the sexes were equally divided.

Testing Instruments

Silent reading comprehension ability was measured by the *Gates-MacGinitie Silent Reading Test*. Oral reading performance was determined by the *Gilmore Oral Reading Test Form C* (1968). The *Lorge-Thorndike Intelligence Test, Level 2, Form A* was administered to all the subjects. The auditory memory span subtests of the *Weschler Intelligence Scale for Children* (1949), and the visual memory span subtest of the *Detroit Test of Learning Aptitudes* (1967) were used to measure memory span. A Word Recognition Test was composed by the investigator and included all the words occurring at least once in the *Gilmore Oral Reading Test, Form C*.

The investigator's copy of the *Gilmore Oral Reading Test, Form C* was divided into syntactic constituents, using the linguistic algorithm proposed by Latham (1972) to determine syntactic constituents (see Appendix B). The *Esterline Angus Speedservo AZAZ portable Labgraph* recorded on a strip chart, a graphic display of the speech sounds, as oscillations to the right of the baseline on the ten-point amplitude scale of the chart, while pauses were displayed as time intervals along the baseline (see Appendix A).

Statistical Analysis

The basic design of the study was a three-way Analysis of Variance over silent reading comprehension groups, on a number of pause measurements obtained from the oral reading performance of the test sample. These pause measurements constituted the main dependent variables in the study and were obtained objectively. Other dependent variables included auditory and visual memory span measurements, and intelligence.

In addition, the relationship between the oral reading comprehension scores of the test sample, and the various dependent variables included in the study, were examined by means of correlation coefficients and a three-way Analysis of Variance, grouping by oral reading comprehension, grade and sex.

Data were analyzed over the entire oral reading material read by each subject, and also over the first 70 syntactic constituents, which constituted the test material read in common by the entire sample.

Main Findings and Conclusions

The percentage of reading time spent pausing while reading orally was able to discriminate between silent reading comprehension groups of varying degrees of ability. This finding tends to lend support to the contention

of Pival (1971), Goodman (1968) and McCracken (1967) that the silent and oral reading processes of young children may be very similar.

Although the time spent pausing while reading orally was able to discriminate significantly between silent reading comprehension groups, all the children tended to resist a disruption within a syntactic constituent if at all possible. This finding agrees with the research on the processing of oral language, especially that of Wilkes and Kennedy (1970), Martin, Kolodziej, and Genay (1971), and Brown (1971), even though their research was done with adult readers.

The number of pauses the subjects made within syntactic constituents, however, seemed to depend more on the semantic and syntactic difficulty of the reading material than on their word recognition abilities. This finding seemed to support Goodman and Niles (1970) statement that the perceptual, syntactic and semantic information used in the reading process are used simultaneously and not sequentially; and also Brown's (1971) findings that 64 per cent of pause variance in oral reading can be predicted from syntactic measures.

Although all the pausing variables were able, in one way or another, to discriminate between silent reading comprehension scores, there was no statistical evidence in this study to indicate any relationship at all between oral reading comprehension scores and any of the pausing variables measured. Further, the data revealed no relationship between oral reading comprehension scores and silent reading comprehension scores.

Since the data do indicate that the use of the pause phenomena in oral reading does differentiate significantly between silent reading comprehension groups, it seems that the model of silent reading shown in Figure 1 appears to be correct, and that there actually may be an element of aural language input in silent reading which assists in comprehension, even when reading silently.

However, when the child is asked to read orally, it appears from these data that there is no relationship between how he is processing his visual input (i.e. syntactic constituents), and the meaning (or product) he obtains from the passages.

An examination of Figure 2 may provide two possible explanations for this occurrence: (1) in the process of recoding to aural input-oral output and then decoding to meaning, the child is unable to finish the process by recalling what he has "heard", either because he cannot remember or because he neglected to recode to aural input. He recoded directly to oral output from graphic input and did not "listen" to what he was reading; or (2) he stops the process at oral output and does not even attempt to complete processing for meaning. He is therefore only able to grasp what little comprehension of the material he has obtained from the element of direct decoding from graphic input to meaning, as shown by the broken line in the model. This could account for the lack of relationship between how the child seems to be processing his visual input while reading orally and oral reading comprehension, while at the same time the actual silent and oral reading processes may be very similar (see Figure 1 and Figure 2).

A perusal of the individual oral reading comprehension scores, however, did indicate that the younger and less proficient silent readers did seem to

be better able to "listen" to themselves while reading orally, probably because they are still using a strong auditory component while reading silently also. Consequently, the auditory "feedback" from an oral reading performance tended to help their comprehension of the passage.

There was a lack of significant relationship between memory span (both visual and auditory) and the processing of written language in this study. However, a closer inspection of the data suggested that the relationship between the processing of written language and memory span is something that develops, or is learned, but was not really evident in this study until towards the end of the grade three level.

Although intelligence did appear to have some relationship to how a child was using the pause while reading orally, the relationship was closely tied in with the age of the child, and the syntactic and semantic difficulty of the material which he was reading.

The data also indicated that semantic and syntactic difficulty of the reading material possibly has more bearing on the ability of the children to comprehend the material than does their ability to recognize the words.

Significant contributions to the effectiveness of reading, arranged in order of rank as determined by the data in this study seem to be: the efficiency of the reading process used by the child, the syntax and semantics of the material to be read, and lastly, the word recognition skills of the reader.

Implications of the Study

The implications of this study were divided into three sections: implications for (1) reading theory, (2) reading research, and (3) the teaching of reading.

Reading Theory

The data indicate that the syntactic constituent, as defined in this study, is very probably a functional linguistic and cognitive unit in the perception of written language.

Further, timing cues (or pauses) in oral reading, seem to be able to differentiate between the silent reading comprehension abilities of children who have been exposed to reading instruction for not more than 2.5 years, and also provide some indication of how these children may be processing written language.

The products of these processes (or the comprehension of the text), however, do not appear similar in the two situations—silent reading and oral reading. It seems then that although an observation of where pauses occur in oral reading output may be a good indication of the reading "processes" employed by young children, silent reading is a better instrument to use if the "product" (or comprehension) is to be observed.

There also seems to be a close relationship between the surface structure of the text, the number and length of pauses made while reading orally, and the silent reading comprehension ability of the reader. Pausing in oral reading might then be used to contribute to insights into whether a reader is able to simultaneously process the syntactic and semantic content of a passage, which may be the "linguistic sophistication" referred to by Chomsky and Halle (1968, p. 50) and which they contend is necessary for acceptable oral reading.

All the data reemphasized that the ability to process written language is a developmental process. At the grade three level there was sufficient evidence to indicate a very important developmental stage, and several factors occurred that were not present at the grade two level.

Reading Research

The pause (or juncture) is only one aspect of the entire suprasegmental system. Perhaps further research, based on this study and using the same, or similar instrument for obtaining objective measurements, could be initiated, using the suprasegmental phonemes of pitch and/or stress as the dependent variable(s).

An attempt might also be made to study the actual use of the pause in silent reading, and silent reading processes, by using the newest type of eye-camera (the oculometer, which is attached to a computer system), and which is able to locate a visual fixation of a given length of time, at a precise point in the text.

Much more research needs to be done in the area of memory span and the processing of written language by young children.

Teaching Reading

The findings of this study emphasize the importance of the language component in teaching reading—that is, especially the syntax of the language, rather than the recoding aspect of reading, which stresses word recognition skills. This study has shown that children who scored low on silent reading comprehension tended to combine syntax (or syntactic constituents) and timing cues (pauses), which were inappropriate to each other. In addition, the data also indicated that adequate comprehension in silent reading was possible without thorough word recognition skills. These data suggest strongly what Goodman (1968) and Weber (1968) have already proposed—that more emphasis should be placed on the syntactic structure of the reading text, rather than on words as such.

The investigator also feels that although oral reading may be used as a diagnostic tool, nevertheless, it may not be an accurate indication of how well a child is able to comprehend silent reading material. Since comprehension is the ultimate goal of reading, then oral reading does not appear to have any merits in attempting to teach or to test comprehension.

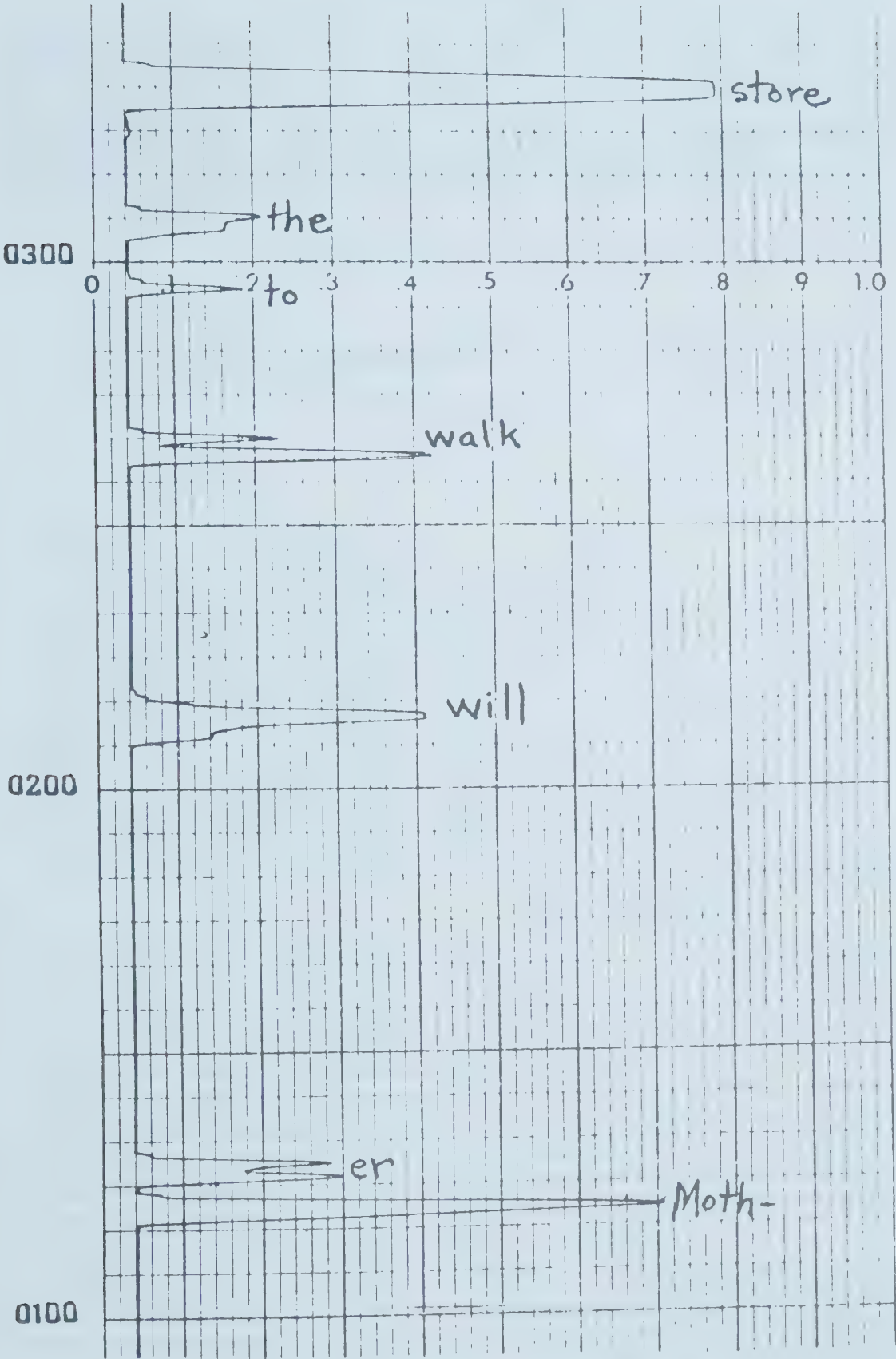
Finally, the children who scored lowest on comprehension in this study seemed to possess the best rote memories. It may be that these children tend to rely too specifically on this type of learning. Teachers should be aware of this, and test comprehension of reading material by requiring responses that do not utilize rote memory.

References

- Bond, Z. S. Units of speech perception. In *Working papers in linguistics*, No. 9. Computer and Information Science Research Centre, Ohio State University, Columbus, Ohio, 1971, 1-112.
- Brown, E. R. Lexical and syntactic predictors of the distribution of pause time in reading, (Doctoral dissertation, Syracuse University.) Ann Arbor, Mich.: University Microfilms, 1971, No. 71-18468.
- Chomsky, N., & Halle, M. *The sound pattern of English*. New York: Harper & Row, 1968.

- Cromer, W. The difference model: A new explanation for some reading difficulties. *Journal of Educational Psychology*, 1970, 61 (6), 471-484.
- Dale, H. C. A. Retroactive interference in short-term memory. *Nature*, 1964, 203, 1408.
- Fodor, J. A., & Bever, T. G. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 414-420.
- Garret, M., Fodor, J. A., & Bever, T. The active use of grammar in speech perception. *Perception and Psychophysics*, 1966, 1, 30-32.
- Goldman-Eisler, F. *Psycholinguistics—Experiments in spontaneous speech*. New York: Academic Press, 1968.
- Goodman, K. S. (Ed.) *The psycholinguistic nature of the reading process*. Detroit: Wayne State University Press, 1968.
- Goodman, K. S., & Niles, O. S. *Reading process and program*. Urbana, Ill.: N. C. T. E. Publication, 1970.
- Kolers, P. A. Three stages of reading. In H. Levin & J. Williams (Eds.), *Basic Studies on reading*. New York: Basic Books, 1970.
- Latham, R. O. M. The relationship of patterns of written language segmentation to reading comprehension. A dissertation proposal submitted to the Department of Elementary Education, University of Alberta, 1971.
- Latham, R. O. M. Personal communications, "On the syntactic constituent," January 7, 1972; March 17, 1972; July 11, 1972.
- McCracken, R. A. The informal reading inventory as a means of improving instruction. *Perspectives in reading, the evaluation of children's reading achievement*, 1967, 8, 85.
- Martin, J. E., Kolodziej, B., & Genay, J. Segmentation of sentences into phonological phrases as a function of constituent length. *Journal of Verbal Learning and Verbal Behavior*, 1971, 10 (3), 226-233.
- Miller, G. A., & Isard, S. Some perceptual consequences of linguistic rules. *Journal of Verbal Learning and Verbal Behavior*, 1963, 2, 217-228.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Crofts, 1967.
- Neisser, U. The role of rhythm in active verbal memory serial intrusions. *American Journal of Psychology*. 1969, 82 (4), 540-546.
- North, A. J., & Jenkins, L. B. Reading speed and comprehension as a function of typography. *Journal of Applied Psychology*, 1951, 35, 225-228.
- O'Connell, D. C., Turner, E. A., & Onuska, A. Intonation, grammatical structure, and contextual association of immediate recall. *Journal of Verbal Learning and Verbal Behavior*, 1968, 7, 110-116.
- Osgood, C. E., & Sebeok, J. A. (Eds.) *Psycholinguistics*. *Journal of Abnormal Psychology*, 1954, 49 (4, Pt. 2), 8-49.
- Pival, J. G. Stress, pitch and juncture: Tools in the diagnosis and treatment of reading ills. In E. Dechant (Ed.), *Detection and correction of reading difficulties*. New York: Appleton-Century-Crofts, 1971.
- Smith, F. *Understanding reading: A psychological analysis of reading and learning to read*. New York: Holt, Rinehart & Winston, 1971.
- VanUden, A. M. J. *A world of language for deaf children. Part 1: Basic principles*. Rotterdam: Rotterdam University Press, 1970.
- Weber, R. M. The study of oral reading errors: A survey of the literature. *Reading Research Quarterly*, 1968-69, 4, 96-119.
- Wickelgren, W. A. Acoustic similarity and retroactive interference in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1965, 4, 53-61.
- Wickelgren, W. A. Context-sensitive coding, associative memory, and serial order in speech behavior. *Psychological Review*, 1969, 76, 1-15.
- Wilkes, A. L., & Kennedy, R. A. The relative accessibility of list items within different pause-defined groups. *Journal of Verbal Learning and Verbal Behavior*, 1970, 9, 197-201.

APPENDIX A

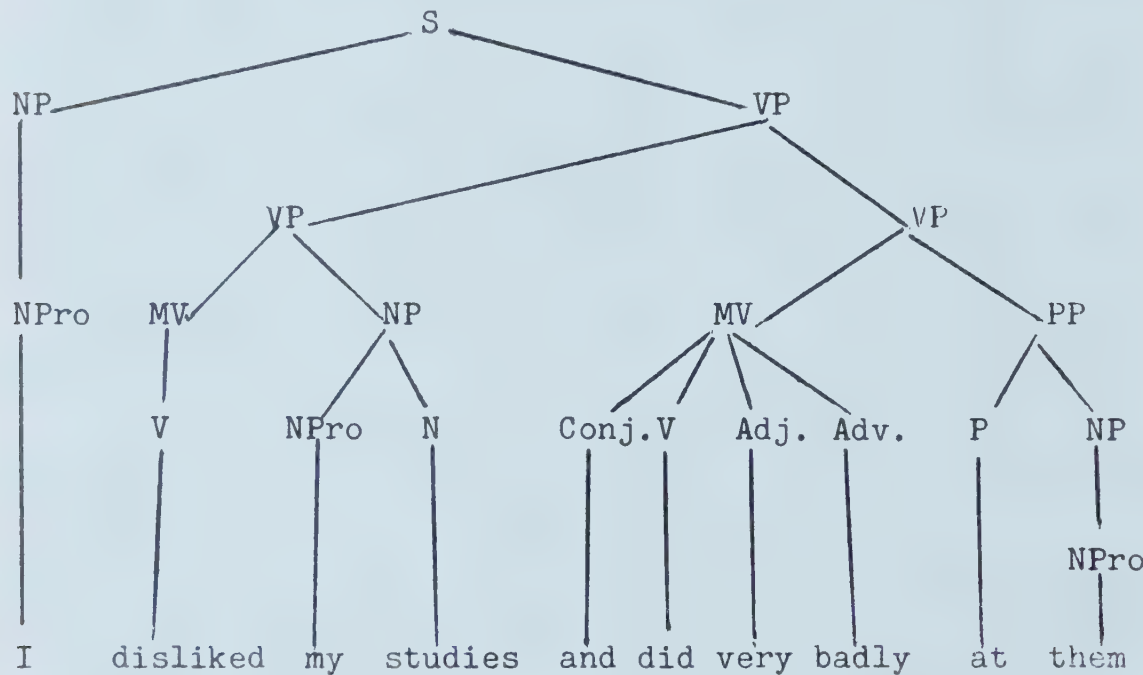


GRADE TWO BELOW-AVERAGE READER

APPENDIX B
SYNTACTIC CONSTITUENT

The term “Syntactic Constituent” will refer to the lowest major constituent as defined by Latham (1971). It may be formed from all labelled nodes in a surface structure tree. These constituents are found by locating those nodes which are immediately above the lexical nodes and deciding whether or not there are sister-nodes to the lexical nodes. If there are no sister-nodes to a specific lexical node, then the node immediately dominating that lexical node is a lowest major constituent. If there are sister nodes to a specific lexical node, then the lowest major constituent associated with that node is the node which immediately dominates all sisters of the lexical node in question.

Example: I disliked my studies and did very badly at them.



Lexical nodes: I dislike my studies and did very badly at them

Sister nodes: (1) my studies

(2) and did very badly

The lowest major constituent associated with the sister-nodes is the NP node for (1) and the MV node for (2)

G. BROADLEY

and

KATHLEEN M. BROADLEY

County of Newell

Rural Standardisation of the Burt-Vernon Graded Word Reading Test

A rural standardisation conducted in the County of Newell, 120 miles from Calgary, confirmed the generalisability of the Burt-Vernon Graded Word Reading Test. Norms closely paralleled those established by Vernon in Calgary in 1973. A significant difference was found among younger children, however, indicating that the Calgary norms may be pro-rated too far downwards. It would appear also that with the use of a phonetically based programme, the rate of learning sight words accelerates rapidly once particular skills have been mastered. An hypothesis that girls taught by women teachers would perform significantly better than boys was not upheld. (Guy Broadley and Kathleen Broadley are with the Education Department of the County of Newell in Brooks, Alberta.)

This study was a replication of Vernon's re-standardisation of the Burt-Vernon Graded Word Reading Test (Vernon, 1973). It was conducted in the County of Newell No. 4 where the total school population is in the vicinity of 1,500 pupils. There are eight public schools in the County which covers an area of 2,350 square miles. The purpose of the replication was to compare rural Albertan reading attainments with those of Calgary children, and to provide local norms on a graded word reading test. Although other word reading tests are available, the Calgary restandardisation was being used in the County. It was felt that the norms of this test would be more accurate in an area the central point of which is only 120 miles from Calgary.

The test itself is widely used to establish word reading attainment, and is useful in diagnosing the types of errors that are involved in the child's decoding, even though its prime purpose is not diagnostic. The test is easily and quickly administered and the results, coupled with the examiner's clinical observations, often provide direction for further diagnosis of reading difficulty.

When Vernon restandardised the test in 1973, he found that the Calgary children of that year learned later than the Burt sample of 50 years before, but that the rate of learning, once begun, was faster. This was apparent in the findings until the age of 12½ was reached. At that age convergence took place. Factors such as later school entry, less emphasis on oral reading, and higher socioeconomic status were mentioned by Vernon as possible contributors to the differences found. A change in the frequencies of words recognised by the Calgary sample was reported, and a revised order of the test words was recommended for local use.

Due to the relatively standard form of curriculum and instruction in the rural and urban areas of Alberta, the authors hypothesised that overall, little difference would be found between the Calgary norms and those that would be developed from the rural sample. However, it was expected that after initial contact with formal reading instruction, when the child learned to read a few words and understand the simpler mechanics of decoding, a rapid increase in reading development would take place and that this would be reflected in the scores. It was hypothesised that Vernon's norms, through being downward pro-rated, would not reflect this phenomenon.

Vernon (1973) suggested that exposure to media may be a factor that has contributed to the establishment of the order of knowledge of words. As rural children are now exposed to much the same media as city children, it was hypothesised that the frequency of recognition of individual words in the test would be very similar to the Calgary findings.

There is already considerable evidence suggesting that boys lag behind girls in reading development, at least until the end of elementary school. Kagan (1964) suggested that the discrimination of observed sex-role activities contributed much to the difference between the reading attainments of boys and girls. Preston (1962) noted that the German boys of that time scored better in reading than the girls and that German elementary teachers were predominantly male. As all teachers but one of grades four and below within the County sample were female, it was hypothesised that girls would attain better than boys up to age ten.

Methodology

Sampling

The total population, grades one to nine, in the eight schools of the County of Newell was ascertained from school returns for October 31, 1974. The only schools in the County that were excluded were the three Hutterite schools. It was decided by the authors that 180 children would provide a manageable sample. This was approximately one child in seven of the grade one to nine population. The schools were randomly ordered and numbered. Children in the respective grades within their own schools were alphabetized regardless of sex. A number less than eight was drawn and this decided the first member of the sample, in grade one, in the first school. Every seventh child was selected thereafter in that and the succeeding grade one classes in the sample, then on through the other grades, selecting from all the schools in order at each grade level.

If a member of the sample was absent when the testing was taking place, the child immediately alphabetically prior was taken into the sample as a sub-

stitute. If this child was also absent the one immediately after the original choice was included. In these substitutions no other criteria were taken into account.

Four students were later dropped from the sample because they were the only 15-year-old children tested. They were not regarded as representative of their age group as most 15-year-old children are in grade ten. The remaining 170 students made up the final sample. They were evenly distributed in grades one through nine with the lowest number in any grade being 17, and the largest, 22. The age range was from 6 years 0 months to 14 years 10 months. There was an equal distribution of boys and girls.

Data Collection

The data were collected in one week at the end of the first school term of the 1974-75 school year. The testing was done by the authors in the following way:

The respondent was shown the printed form with the test on it (Calgary revised order was used) and asked to read each word. Responses were recorded on a duplicate form. After five consecutive errors the respondent was encouraged to look ahead in the test for other words that might be attempted. If a response on any word was not begun after fifteen seconds the child was asked to proceed to the next word.

Thus, test administration was a little different from standard procedure, but as the child still had the opportunity to make an attempt on every word, it was felt that the slightly streamlined administration procedure employed would not affect scores and would speed up the testing.

The respondent's age, sex, grade, and score (number of correctly pronounced words) were recorded on the form. To determine the reliability of the scores, the standard errors of the mean scores of three groups containing successive age groups were found. It seemed appropriate to include at least the levels immediately above and below because given chronological ages will incorporate several reading ages (and vice versa). Standard errors thus arrived at were 2.6 words (6-8 years), 2.2 words (9-11 years), and 1.7 words (12-14 years). The 5% confidence limits ranged from five words at the lower age levels to three words among the older children. The authors have integrated these standard errors into a norm structure that has converted the decimals of years into months, making comparisons with other test results a little simpler. This is available on request from the authors.

Norming

A scattergram was drawn showing age in years and months on the x axis, and score on the test on the y axis. The median age of children who obtained each score was computed. These median ages were joined on the scattergram and the curve smoothed. This is the same method of norming as used by Vernon in the 1973 Calgary restandardisation. A smoothed curve through the median ages of each group of ten scores (0-9, 10-19, etc.) proved to be identical.

Discussion of Results

The hypothesis that little difference would be found between the norms of the urban and rural samples was upheld. The two sets of scores were compared at the various median ages (reading ages). Only below median age 7½ were the

TABLE 1
A COMPARISON OF THE URBAN (VERNON, CALGARY) AND
RURAL (COUNTY OF NEWELL) NORMS OBTAINED ON THE BURT-VERNON
GRADED WORD READING TEST

Calgary 1973		County 1974		Calgary 1973		County 1974		Calgary 1973		County 1974	
R.A. ^a	Score ^b	R.A.		R.A.	Score	R.A.		R.A.	Score	R.A.	
6.3	1	6.7		7.9	37	7.8		10.8	73	10.6	
6.3	2	6.7		8.0	38	7.8		10.9	74	10.7	
6.3	3	6.7		8.1	39	7.9		11.0	75	10.8	
6.3	4	6.7		8.1	40	8.0		11.0	76	10.9	
6.3	5	6.7		8.2	41	8.1		11.1	77	11.0	
6.4	6	6.8		8.3	42	8.2		11.2	78	11.2	
6.4	7	6.8		8.3	43	8.3		11.3	79	11.3	
6.5	8	6.8		8.4	44	8.3		11.4	80	11.5	
6.5	9	6.9		8.5	45	8.4		11.5	81	11.6	
6.5	10	6.9		8.6	46	8.4		11.6	82	11.8	
6.6	11	6.9		8.7	47	8.5		11.7	83	12.0	
6.6	12	7.0		8.7	48	8.5		11.8	84	12.1	
6.6	13	7.0		8.8	49	8.6		11.9	85	12.3	
6.7	14	7.0		8.9	50	8.7		12.0	86	12.5	
6.7	15	7.1		9.0	51	8.8		12.1	87	12.6	
6.7	16	7.1		9.1	52	8.9		12.2	88	12.7	
6.8	17	7.1		9.2	53	8.9		12.3	89	12.8	
6.8	18	7.2		9.3	54	9.0		12.4	90	13.0	
6.9	19	7.2		9.3	55	9.1		12.5	91	13.1	
6.9	20	7.2		9.4	56	9.1		12.6	92	13.2	
7.0	21	7.2		9.5	57	9.2		12.7	93	13.3	
7.0	22	7.2		9.6	58	9.3		12.8	94	13.4	
7.1	23	7.2		9.7	59	9.4		13.0	95	13.4	
7.1	24	7.3		9.8	60	9.5		13.3	96	13.5	
7.2	25	7.3		9.8	61	9.6		13.5	97	13.6	
7.2	26	7.3		9.9	62	9.7		13.7	98	13.7	
7.3	27	7.3		10.0	63	9.8		14.0	99	13.8	
7.3	28	7.4		10.1	64	9.8		14.4	100	13.9	
7.4	29	7.4		10.2	65	9.9		14.7	101	14.0	
7.5	30	7.5		10.3	66	10.0		14.9	102	14.3	
7.5	31	7.5		10.3	67	10.1		15.2	103	14.6	
7.6	32	7.5		10.4	68	10.2		15.5	104	14.9	
7.7	33	7.6		10.5	69	10.2		15.8	105		
7.7	34	7.6		10.6	70	10.3		16.2	106		
7.8	35	7.7		10.6	71	10.4		16.5	107		
7.8	36	7.7		10.7	72	10.5		16.8	108		
									109		
									110		

^a Reading Age

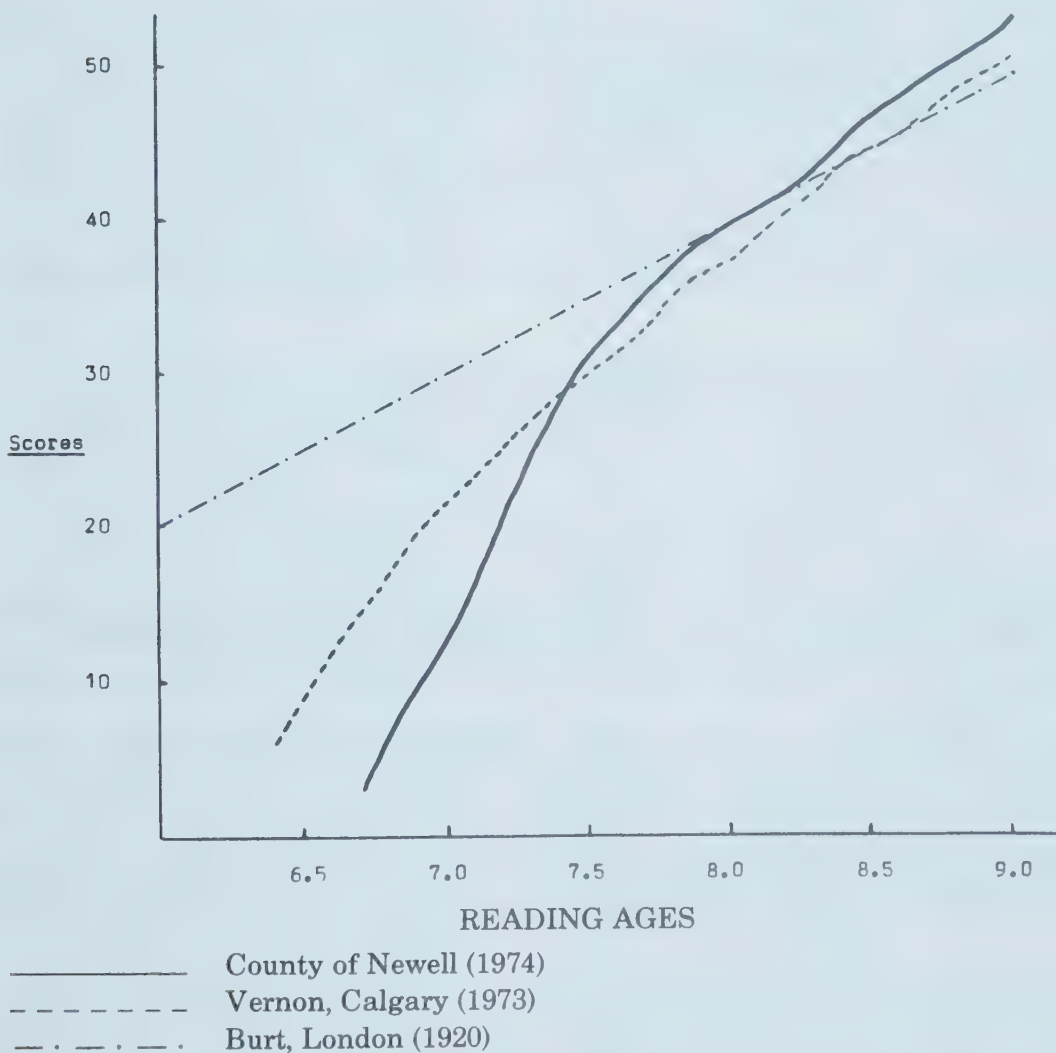
^b The score is the number of words read correctly on the test.

scores obtained markedly different. Here, the mean scores at each median age were compared from median age 6.7, the lowest in the rural sample, through to 7.5. The difference between the means of the two sets of scores at this level was significant ($t = 3.76, 16df, p < 0.01$). This lends support to the second hypothesis that Vernon's downward pro-rated norms, necessitated by later sampling, would not reflect the actual learning curve. This curve appears to demonstrate late but rapid development in the acquisition of a sight vocabulary (see Figure I). The children in the County sample did not begin to recognise words on the form until, on average, half way through year 6.

The emphasis on the phonetic structures in the grade one reading programme develops skills that the children can then apply to many words. The authors contend that the rapid increase develops in much the same way as the language explosion in very young children. Language development of

young children expands suddenly with the use of pivot words. A similar generalisation is applied by children who at first recognise a few words by sight, but then quickly expand their sight vocabulary with the identification and use of simple phonetic elements. Using this learning technique, it would seem that the students in the County sample do not begin word recognition until much later than Burt's London children of 1920, but then learn much faster, an observation made also by Vernon of his Calgary sample in 1973. Further investigation by the authors of the rate of learning sight words at this level is presently being undertaken.

FIGURE I
GRAPH COMPARING THE NORMS OBTAINED FROM THE LONDON (1920),
CALGARY (1973) AND COUNTY OF NEWELL (1974) SAMPLES,
READING AGES 6.0-9.0



The frequency of recognition of individual words was similar to that found by Vernon in 1973. Since the movement of one word is able to alter the position of several others, changes were expected. But, even with no allowances made for any change, among the first twenty words where changes seemed greatest, they were not at all significant. Two later words, LUNCHEON and TONGUE, moved 17 places further down the list. PHILOSOPHER, BINOCULAR, and

UNIQUE, appeared to be recognised more easily by the County sample, each moving 12 words or more up the list. As children are encouraged to at least look at all the words, the word order serves only to facilitate test administration.

When a comparison of boys' and girls' scores was made at the lower level of the test it was found that 10 of the 14 boys in the sample younger than age 7 years 6 months read less than ten words. Of the 12 girls in the same age group only four read less than ten words. Of the total sample, 15 out of the 85 boys read less than 30 words but only 8 out of the 85 girls read less than 30 words. Thirty words is equivalent to reading age 7.5 years. However, the difference in attainments between the boys and girls up to age 11 years was not significant. Nor was the difference up to age 7.5 years significant beyond the 0.2 level. Thus the results do not therefore establish a positive trend, but neither do they completely assuage suspicions.

Aside from the difference found at the younger age levels, the rural sample closely paralleled the urban one of 1973 in attainment on the test and in frequency of recognition of words. While no pretence is made as to the validity of Vernon's norms of 1973 outside of Alberta, there certainly appears to be a likelihood that they will be little different within Albertan school systems where children speak English, start school at six, watch television, and experience similar reading programmes. If it could be assumed that these four variables are as important as they appear, then the test's norms may have even wider validity.

The co-operation of school staffs in the County of Newell #4 is acknowledged.

References

- Asher, S. R., & Markell, R. A. Sex differences in comprehension of high- and low-interest reading materials. *Journal of Educational Psychology*, 1974, 66, 680-687.
- Braine, M. D. S. The ontogeny of English phrase structure: The first phase. *Language*, 1963, 39, 1-13.
- Kagan, J. The child's sex role classification of school objects. *Child Development*, 1964, 35, 1051-1056.
- Preston, R. C. Reading achievement of German and American children. *School and Society*, 1962, 90, 350-354.
- Vernon, P. E. A restandardisation of the Burt-Vernon Graded Word Reading Test. *Western Psychologist*, 1973, 4, 72-77.

L. E. JURY

Central York Schools

D. J. WILLOWER

and

W. J. DeLACY

The Pennsylvania State University

Teacher Self-Actualization and Pupil Control Ideology

The prediction that level of teacher self-actualization would be directly related to humanism in teacher pupil control ideology was supported by data gathered from a sample of 272 public school teachers. Expected differences in the pupil control ideology of elementary and secondary teachers were found, but level of self-actualization was not significantly different for teachers grouped dichotomously according to sex, grade levels taught, or experience. (Dr. Jury is assistant superintendent, Central York Schools, York, Pennsylvania; Dr. Willower and Dr. DeLacy are professors of education, Division of Education Policy Studies, The Pennsylvania State University.)

The concept of self-actualization, discussed earlier by other writers (cf. Goldstein, 1940), but developed chiefly by Maslow (1957, 1964, 1968, 1970, 1971) has been the object of considerable theoretical attention. However, empirical inquiries have been less common. This article reports on an investigation of the relationship between public school teachers' levels of self-actualization and their orientations toward pupil control.

We were interested in probing the relationship between teacher self-actualization and attitudes toward pupil control or discipline in part because control represents an aspect of teachers' work that earlier studies show to be of cardinal importance in school organizations. Concern with pupil control has been characterized as a salient feature of life in schools, affecting teachers' relationships with peers and other colleagues such as counselors and administrators (Willower & Jones, 1967; Willower, 1971). Further, good control is often equated with good teaching (Gordon, 1957), and it has been found that more rigid views regarding discipline are one consequence of teacher socialization (Hoy, 1968). Given social system

pressures for relatively unyielding attitudes on pupil control, the question of whether teacher level of self-actualization would be a predictor of pupil ideology seemed open enough to warrant investigation.

Conceptual Framework

Compared with a person lacking self-actualization, the self-actualizing individual has been conceptualized as one who uses his or her talents and capabilities more fully, functions relatively autonomously, tends to live in the present, and has a more benevolent outlook on human nature and on life (Maslow, 1970). The Personal Orientation Inventory (POI) developed by Shostrom (1968) serves as an operational definition of self-actualization.

Pupil control ideology has been conceptualized on a humanistic-custodial continuum. A humanistic ideology emphasizes an accepting, trustful view of pupils and optimism concerning their ability to be self-disciplined and responsible. A custodial orientation stresses the maintenance of order, distrust of pupils, and a moralistic approach to pupil control. The Pupil Control Ideology (PCI) Form (Willower, Eidell, & Hoy, 1967) furnishes an operational definition for this variable.

To examine the relationship between teacher self-actualization and pupil control ideology, the following hypothesis was proposed: level of self-actualization will be directly related to humanism in pupil control ideology.

The rationale for this hypothesis was relatively straightforward. In contrast with the less self-actualizing person, the self-actualizing individual is characterized by satisfied lower order needs and perceptions of a relatively benign environment. Hence, the self-actualizing teacher is not likely to see students as part of a threatening environment and will not view strict control of students as essential.

Further, the perceived norms of the teacher group may be especially important for the less self-actualized teacher. It has been reported that teachers tend to believe that other teachers hold comparatively custodial views on pupil control (Packard & Willower, 1972). Thus, for the non-self-actualizing teacher, the colleague group could be seen as a source of pressures for custodial orientations toward students.

Procedures

Instruments

The Personal Orientation Inventory (Shostrom, 1968) was developed as a result of a "need for a diagnostic instrument which gives . . . a measure of current level of positive health or self-actualization" (Shostrom, 1964, p. 207). It consists of 150 two-choice items selected from observed value judgments of clinically healthy and clinically troubled persons, and from theoretical formulations of writers in psychology. Although the POI has several scales and subscales, a single combined score provides an overall measure of self-actualization (Shostrom, 1968; Damm, 1972).

Shostrom (1968) reported test-retest reliabilities of .91 and .93 for the POI; claims for its validity were based upon the instrument's ability to discriminate between individuals judged by clinical psychologists to be relatively self-actualized and non-self-actualized. The highest score

theoretically possible on the POI is 150; the higher the score the more self-actualizing the respondent.

The Pupil Control Ideology Form measures educators' orientations toward discipline on a continuum having ideal-type poles represented by humanism and custodialism. The instrument consists of 20 items on a 5-point scale with response categories from "strongly agree" to "strongly disagree." The scoring range is 20-100; the higher the score, the more custodial the ideology. Split-half reliabilities ranged from .91 to .95; the instrument was shown to differentiate between teachers and schools judged to be custodial or humanistic (Willower, Eidell, & Hoy, 1967).

Sample

The instruments were administered to all of the teachers in a single school district in a northeastern state in the fall of 1972. The selected district combined suburban and rural features but bordered a middle-sized city. It had seven elementary schools, a junior and a senior high school. Faculty salary, years of teaching experience, and educational level approximated state averages. The faculty of the district was composed of 137 elementary and 141 secondary teachers.

Findings

Usable returns were secured from a total of 272 teachers, 134 at the elementary level and 138 at the secondary level. The high rate of return, 98 percent, probably resulted because the study was endorsed by both the teachers' association and the school superintendent, and the forms were completed during faculty meetings.

The main hypothesis was tested for the overall sample and separately for elementary and secondary teachers. For these tests, the Pearson product-moment correlation was the statistic employed. These and all other calculations were performed using programs of the Computation Center of The Pennsylvania State University.

Relevant data appear in Table 1. Negative correlation coefficients are a function of the scoring of the instruments. Each of the three tests of the hypothesis produced a significant association between self-actualization and pupil control ideology. Hence, the hypothesis that teacher level of self-

TABLE 1
TEACHER SELF-ACTUALIZATION AND PUPIL CONTROL IDEOLOGY

Teachers	N	SA		PCI		r
		\bar{X}	S.D.	\bar{X}	S.D.	
All	272	96.8	13.1	54.3	10.6	-.39**
Elem.	134	96.6	13.4	50.1	9.7	-.51**
Sec.	138	96.9	12.8	58.3	9.6	-.34*

* $p < .01$

** $p < .001$

actualization will be directly related to humanism in pupil control ideology cannot be rejected.

Because self-actualization (SA) has a developmental aspect and is commonly conceived as a stage in life when a mature personality has been formed, a partial correlation also was computed between SA and PCI for the overall sample, controlling for the effect of age. The partial correlation was $-.37$, significant beyond the $.001$ level. The correlation coefficient for SA and age was $-.16$, quite low but significant at the $.05$ level. This indicates a slight negative association between SA and age for the present sample of teachers.

Other demographic characteristics on which data were gathered included teacher sex, grade level taught (elementary vs. secondary), and experience (more than 5 years vs. 5 years or less). When *t*-tests were computed to examine mean differences on SA for teachers grouped according to these dichotomous variables, none were found to be significant.

On the other hand, significant differences in PCI were found on all three variables. Elementary teachers were more humanistic than their secondary school counterparts; females were more humanistic than males; and less experienced teachers were more humanistic than more experienced ones. See Jury (1973) for details. These findings are consistent with the results of previous investigations (Willower, Eidell, & Hoy, 1973).

Discussion

As predicted, teacher level of self-actualization was significantly associated with pupil control ideology: the higher the level of self-actualization, the more humanistic the ideology. The association between the two variables was strongest for elementary school teachers. That is not surprising since, in elementary schools, norms for pupil control in the teacher collectivity lay less stress on custodial views and behavior than those prevalent in secondary schools.

In this connection, it might be conjectured that more self-actualizing teachers would have fewer peer conflicts and greater work satisfaction in the elementary school than in the secondary school, and hence might be found in greater numbers in elementary schools. Yet self-actualization mean scores did not differ significantly across various teacher groupings. If only those teachers who scored very high, 112 and above, on self-actualization are considered, the distribution is quite even between elementary and secondary levels. Of these 35 self-actualizing teachers, 17 were in elementary schools and 18 were in secondary schools. The elementary school apparently does not attract and retain teachers who have higher levels of self-actualization than those in secondary school.

There have been a substantial number of inquiries on pupil control. Some of them probe relationships that stem from the character of the organizational setting and the teacher group. In general, it seems that social system factors create pressures in custodial rather than in humanistic directions.

Other studies deal with variables that are more individual and psychological in orientation. Thus, there is evidence that, contrasted with more custodial colleagues, teachers who are humanistic in pupil control ideology

tend to be less dogmatic (Willower, Eidell, & Hoy, 1967), exhibit less status obeisance or deference toward authority (Helsel, 1971a), hold emergent rather than traditional values (Helsel, 1971b), have a greater sense of power over school matters (Zeilei, 1971), and score higher on measures of creativity (Halpin & Goldenberg, 1973). All of these investigations, except the last, which examined a sample of college students who planned to become teachers, were based upon relatively large samples of teachers.

The portrait of the more humanistic teacher that emerges is quite consistent with the findings of the present research. That research, which adds self-actualization to the overall picture, fleshes out a portrait composed of features that are likely to be seen as desirable. Who wants to be counted against open-mindedness, creativity, and self-actualization?

However, some basic issues remain. One deals with the relation of the teacher and the social setting. Here the organizational character of the public school is critical. Schools are characterized by vague and diverse goals, ambiguous criteria of success, lack of a widely accepted technology of teaching, a mandated relationship with clients who may or may not be willing participants, high population density, stimulus overload especially for teachers, a host of logistical problems, and political vulnerability. In such circumstances, schools like virtually all organizations, maintain structures that serve stabilizing functions. Routines, rules, and the universalistic processing of clientele are illustrative of such structures (see Willower, 1971). All this leads, in the school, to an opposition of the adult and student groups. The adults speak for the organization; the students variously resist, conform, or adapt. Casting our portrait of the more humanistic teacher against these considerations, it should be clear that constraints abound and the possibility of unanticipated negative consequences is substantial.

However, the association between pupil control ideology and pupil control behavior is problematic. This is a second basic issue. Other work shows a significant but moderate correlation between teacher pupil control ideology and behavior (Helsel & Willower, 1974). As might be expected in the social setting discussed, the correspondence is far from perfect. The results of an earlier study (Rexford, Willower, & Lynch, 1972) that showed some differences in classroom behavior between teachers who were quite humanistic and those who were quite custodial in pupil control ideology is perhaps typical of what can be expected. Using Flanders' observation techniques, it was found that the humanistic teachers were more indirect in classroom verbal behavior but did not differ significantly from the custodial teachers in percentage of teacher talk in the classroom. The school's organizational arrangements appear to afford what could be called a circumscribed leeway for teacher behavior.

A final issue concerns the consequences of teacher pupil control ideology and behavior for pupils. Empirical work is limited. Hoy (1973) reported that, in general, the more custodial the pupil control ideology of the school faculty, the more alienated the students. Pritchett (1973) found that custodialism in teacher pupil control behavior was associated with negative attitudes toward school on the part of students.

Such investigations suggest the viability of humanistic approaches to students. However, organizational and social system factors like those dis-

cussed create pressures that can mute or distort humanistic intentions and efforts. It is important to examine variables like self-actualization that can be employed to characterize teachers and that can be related to desirable teacher attitudes and behavior. Yet it is essential to see the whole within the school's social context.

References

- Damm, V. J. Overall measures of self-actualization derived from the personal orientation inventory: A replication and refinement study. *Educational and Psychological Measurement*, 1972, 32, 485-489.
- Goldstein, K. *Human nature in the light of psychopathology*. Cambridge: Harvard University Press, 1940.
- Gordon, C. W. *The social system of the high school*. Glencoe: Free Press, 1957.
- Halpin, G., & Goldenberg, R. Relationships between measures of creativity and pupil control ideology. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, 1973 (mimeographed).
- Helsel, A. R. Status obeisance and pupil control ideology. *Journal of Educational Administration*, 1971, 9, 38-47. (a)
- Helsel, A. R. Value orientation and pupil control ideology of public school educators. *Educational Administration Quarterly*, 1971, 7, 24-33. (b).
- Helsel, A. R., & Willower, D. J. Toward definition and measurement of pupil control behavior. *Journal of Educational Administration*, 1974, 12, 114-123.
- Hoy, W. K. The influence of experience on the beginning teacher. *School Review*, 1968, 76, 312-323.
- Hoy, W. K. Dimensions of student alienation and characteristics of public high schools. *Interchange*, 1973, 3, 38-52.
- Jury, L. E. Teacher self-actualization and pupil control ideology. Unpublished doctoral dissertation. The Pennsylvania State University, 1973.
- Maslow, A. H. Self-actualization and beyond. In J. F. T. Bugental (Ed.), *Challenges of humanistic psychology*. New York: McGraw-Hill, 1957.
- Maslow, A. H. *Religions, values, and peak experiences*. Columbus: Ohio State University Press, 1964.
- Maslow, A. H. *Toward a psychology of being*. (2d ed.), New York: Van Nostrand Reinhold, 1968.
- Maslow, A. H. *Motivation and personality*. (2d ed.), New York: Harper and Row, 1970.
- Maslow, A. H. *The farther reaches of human nature*. New York: Viking Press, 1971.
- Packard, J. S., & Willower, D. J. Pluralistic ignorance and pupil control ideology. *Journal of Educational Administration*, 1972, 10, 78-87.
- Pritchett, W. The relationship between teacher pupil control behavior and student attitudes toward school. Unpublished doctoral dissertation. The Pennsylvania State University, 1973.
- Rexford, G. E., Willower, D. J., & Lynch, P. D. Teachers' pupil control ideology and classroom verbal behavior. *Journal of Experimental Education*, 1972, 40, 78-82.
- Shostrom, E. L. An inventory for the measurement of self-actualization. *Educational and Psychological Measurement*, 1964, 24, 207-218.
- Shostrom, E. L. *Manual for the personal orientation inventory: An inventory for the measurement of self-actualization*. San Diego: Educational and Industrial Testing Service, 1968.
- Willower, D. J. The teacher subculture. In L. W. Drabick (Ed.), *Interpreting education: A sociological approach*. New York: Appleton, Century, Crofts, 1971.
- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology*. University Park: Penn State Studies No. 24, 1967.

- Willower, D. J., Eidell, T. L., & Hoy, W. K. *The school and pupil control ideology* (2d ed. with an annotated bibliography of pupil control studies). University Park: Penn State Studies No. 24, 1973.
- Willower, D. J., & Jones, R. G. Control in an educational organization. In J. D. Raths & others (Eds.), *Studying teaching*. Englewood Cliffs, N. J.: Prentice-Hall, 1967.
- Zelei, R. A. Relationship between pupil control ideology and sense of power of teachers in selected public schools. Unpublished doctoral dissertation. University of Akron, 1971.

The Alberta Journal of Educational Research

Volume XXI, 1975

TABLE OF CONTENTS

- Arikado, M. S. *Status consensus as it relates to team teacher satisfaction*. No. 2, June, 104-109.
- Blackall, B. M. *The works of Jean Piaget published in English*. No. 1, March, 63-70.
- Breiter, J. C. *Reading or listening: A comparison of techniques of instruction in elementary social studies*. No. 2, June, 130-140.
- Broadley, G., and Broadley, K. M. *Rural standardisation of the Burt-Vernon graded word reading test*. No. 4, December, 289-294.
- Card, B. Y. *Community vocational centres in northwestern Alberta*. No. 3, September, 168-182.
- Clarke, S. C. T., and Coutts, H. T. *Toward teacher education in the year 2000*. No. 4, December, 221-240.
- Clifford, M. M. *Validity of expectation—A developmental function*. No. 1, March, 11-17.
- Das, J. P., Manos, J., and Kanungo, R. N. *Performance of Canadian native, black and white children on some cognitive and personality tests*. No. 3, September, 183-195.
- Eagan, Sr. R. *An investigation into the relationship of the pausing phenomena in oral reading and reading comprehension*. No. 4, December, 278-288.
- Eastwood, G. R. *Educational research: Concepts and models*. No. 2, June, 71-83.
- Forester, A. D. *Learning the language of reading—An exploratory study*. No. 1, March, 56-62.
- Foster, L., and Nixon, M. *The interview reassessed*. No. 1, March, 18-22.
- Friesen, D. *A study of a public school system, Hessen, West Germany*. No. 2, June, 84-103.
- Froese, V. *Does the Dolch do?* No. 1, March, 23-27.
- Hobbs, E. D. *Methodological problems in conservation testing with particular reference to volume conservation*. No. 4, December, 262-277.
- Jansson, L. C. *The judgment of simple deductive arguments by pre-service elementary school teachers*. No. 1, March, 1-10.
- Jury, L. E., Willower, D. J., and DeLacy, W. J. *Teacher self-actualization and pupil control ideology*. No. 4, December, 295-301.
- Kehoe, J. W. *Demonstrating the relationship between values and attitudes as a means of changing attitudes*. No. 3, September, 207-212.
- Klein, J. P. *Socratic dialogue vs. behavioural practice in the development of coping skills*. No. 4, December, 255-261.
- Maguire, T. O., Patsula, R. B., and Evanechko, P. O. *The development of word meaning discrimination in children*. No. 3, September, 154-167.
- Martin, J. *Controlled vs. natural setting—Some implications for behavioural analysis and change in classroom situations*. No. 1, March, 39-45.
- McCatty, C. A. *Patterns of learning projects among professional men*. No. 2, June, 116-129.
- Nelson, D., and Sawada, D. *Studying problem solving behavior in young children—Some methodological considerations*. No. 1, March, 28-38.
- Nixon, M., and Gue, L. *Women administrators and women teachers: A comparative study*. No. 3, September, 196-206.
- Ollila, L. O., and Chamberlain, L. A. *The effect of noise and object on acquisition of a sight vocabulary in kindergarten children*. No. 3, September, 213-219.

- O'Reilly, R. *Classroom climate and achievement in secondary school mathematics classes*. No. 4, December, 241-248.
- Poole, H. E. *A logically first stage in training creative teachers?* No. 3, September, 143-153.
- Pritchett, W., and Willower, D. J. *Student perceptions of teacher pupil control behavior and student attitudes toward high school*. No. 2, June, 110-115.
- Robitaille, D. F. *Classroom personality patterns of teachers of secondary mathematics*. No. 4, December, 249-254.
- Verma, S., and Peters, D. L. *Day care teacher practices and beliefs*. No. 1, March, 46-55.

New Zealand Journal of Educational Studies

Published twice yearly by the
New Zealand Council for Educational Research

Editor: P.J. Lawrence

Recent Articles

The Role of Grammar in a Secondary School English Curriculum
W.B. Elley, I.H. Barham, H. Lamb and M. Wylie

Educational Innovation: The Case of the New Zealand Playcentre
G. McDonald

The Educational Philosopher and Educational Research
R.P. Precians

The Development of Morphological Rules in Children with Different
Language Backgrounds
Marie M. Clay

Reduction of Disruption Behaviour in the Classroom: Group and
Individual Reinforcement Contingencies Compared
M.D. Ellery, N.M. Blampied and W.A.M. Black

Attending Behaviour and Reading
C.B.J. Harper and E.N. Graham

Social Participation and Special Class Attendance in
Mildly-Retarded Children
Rosemary F. Smart and Keri M. Wilton

The Effect of Frequent School Changes on Children's Self
Concept Development
G.A. Wagner and G.I. Feletti

Annual Subscription rate: Canada and the USA, US\$8.00
All business correspondence and subscriptions should be addressed to
The Executive Officer, NZCER, Box 3237, Wellington, New Zealand.

CONTEMPORARY EDUCATION

The NOW Journal
of Teaching and Learning

CE: Publishes articles by
teachers, professors and
school administrators

CE: Reviews current books
in the broad field of
education

CE: Helps educators with their
professional tasks

CE: Offers lively editorials

CE: Gives you regular columns
on provocative topics,
people

a. A View from the Bridge—
David Turney*

b. Mostly Miscellany—
M. Dale Baughman**

*Dean, School of Education, Indiana State University

**Professor of Secondary Education, and Editor of CONTEMPORARY EDUCATION

CONTEMPORARY EDUCATION, Reeve Hall,
Room 201-204, Indiana State University,
Terre Haute, Indiana 47809

Please send the next 4 issues of CONTEMPORARY EDUCATION
for \$7.00 (\$10.00 for foreign countries).

_____ Payment enclosed _____ Bill me

Name _____

Address _____

City _____

State _____ Zip _____

METHUEN ANNOUNCE

The British Journal of Teacher Education

Editor: PROFESSOR EDGAR STONES,
University of Liverpool

This new journal, designed to keep all concerned with teacher education abreast of world-wide developments in a rapidly changing field, examines the whole process of teacher education, the policies relating to its development, its institutional patterns, the changing content of initial preparation, in-service education and staff development. Two objectives of the editorial policy of the journal are that it should be actively interested in current issues, and that, with the help of a panel of overseas correspondents, it should reflect developments in teacher education on an international scale.

The journal appears three times a year, in January, April and October.

Subscription rates: Individuals £5.50; Institutions £6.75; North America £6.75 (U.S. \$18.00); airmail to all parts of the world £10.75 (U.S. \$29.00); single copies £2.25 (U.S. \$6.00). All rates inclusive of postage.

Subscriptions to: Methuen and Co., Subscriptions Department, North Way, Andover, Hampshire, SP10 5BE, England.

